METHODOLOGICAL STUDIES OF HEALTH RESEARCH

METHODOLOGICAL STUDIES OF THE HEALTH RESEARCH LITERATURE: CHARACTERIZING NOMENCLATURE, STUDY DESIGNS, AND REPORTING PRACTICES

By DAERIA O. LAWSON, HBSc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2023) Hamilton, Ontario (Health Research Methodology – Clinical Epidemiology Specialization)

TITLE:	Methodological Studies of the Health Research Literature: Characterizing Nomenclature, Study Designs, and Reporting Practices
AUTHOR:	Daeria O. Lawson, Hons. B.Sc. (University of Toronto)
SUPERVISOR:	Dr. Lawrence Mbuagbaw
NUMBER OF PAGES:	xiii, 191

LAY ABSTRACT

Methodological studies are a fundamental component of health research because they evaluate how and why research is done. Over the past four decades, they have been instrumental in uncovering poor practices and have led to improvements in the design, conduct, analysis, and reporting of studies. This field is diverse, and researchers have taken different approaches to report their work. As a result, these studies have been labelled in different ways (e.g., meta-epidemiological study, methodological review, systematic survey) leading to difficulties in searching for and finding them in databases. This thesis outlines a comprehensive evaluation of the landscape of methodological studies in health research with humans by way of: (1) a methodological guidance paper for a type of study—pilot studies—to illustrate how methodologists go about uncovering deficiencies and advising on research practices; (2) a pilot study testing methods to identify methodological studies in literature databases; and (3) a review of methodological studies to characterize this field. Collectively, this work points to the conclusion that streamlining the reporting and labelling of methodological studies is necessary.

ABSTRACT

Background and Objectives:

Methodological studies of health research are undertaken to investigate the practice of research. They have been instrumental in inciting developments in the design, conduct, analysis, and reporting of health research. Due in part to the field's diversity, these studies can be difficult to identify in databases. As these studies have not been comprehensively examined to date, the overarching goal of this thesis was to characterize methodological studies and to investigate how they have been labelled and reported in the literature.

Methods:

First, we demonstrate how methodological studies are conducted to provide guidance to end-users—in this case physiatrists and rehabilitation researchers—in a methods guidance paper on pilot and feasibility studies (PAFS), a type of health research design. Second, we performed a pilot study testing the feasibility of searching for and identifying methodological studies in literature databases. Third, based on the pilot study findings and previous research, we outline a protocol for the development of a reporting guideline for methodological studies of health research. Lastly, as part of the first phase of the reporting guideline development process, we performed a review of methodological studies focusing on those that specifically investigated PAFS.

Results and Conclusions:

In a case study of rehabilitation research, a third of studies labelled as PAFS did not outline any feasibility outcomes, and few provided progression plans to definitive studies.

Guidance was focused on providing recommendations and resources for assessing feasibility to help reduce the prevalence of small studies disguised as PAFS, which wastes research resources. In the pilot of methodological studies, preliminary findings on nomenclature and reporting reinforced the notion that there are many names used to describe studies with similar intentions. It was also determined feasible to build a search strategy to identify methodological studies in literature databases. Subsequent findings from the review of methodological studies illustrated that reporting practices are the most common aspect of research investigated. Study design names such as 'methodological review', 'systematic review', and 'systematic survey' were often used to describe studies with similar motives, i.e., to synthesize data from previously published research, whether the synthesis approach was quantitative or qualitative. Existing reporting checklists were rarely used, and when used not appended, possibly due to irrelevance of fields oriented to studies with persons. This work demonstrates the necessity and importance of consensus on reporting and nomenclature for making methodological studies more accessible to the health research community.

Acknowledgements

I would like to express my utmost gratitude to my supervisor and mentor, Dr. Lawrence Mbuagbaw. Thank you for your generosity, encouragement, trust, and friendship; for supporting my aspirations and training as an epidemiologist; and for sharing with me the many lessons and opportunities to help nurture my professional and personal growth. Most of all, thank you for your dedication and for ensuring my success every step of the way. I strive to emulate and pass on the kind of mentorship you have showed me to others.

To another important mentor during my PhD, Dr. Lehana Thabane, with sincere thanks for always making the time, for your kindness, and for opening many doors for me. I am thankful to my supervisory committee, Dr. Gordon Guyatt and Dr. Behnam Sadeghirad, for challenging me in the best way and for your willingness to guide and see me through to the end of this journey. Your wisdom and stimulating conversations have helped me to become a better and more thoughtful researcher. I am also grateful to the many colleagues and collaborators at McMaster University and beyond, who have helped me and supported this body of work in numerous ways. I am honoured to have stood on the shoulders of giants in the truest sense of the expression.

I would like to thank the funders of my doctoral research including the Ontario Graduate Scholarship from the Province of Ontario and McMaster University, the Ontario Drug Policy Research Network, and the Arthritis Alliance of Canada.

I am incredibly grateful for my mentors, colleagues, and lab mates at the University of Toronto and Toronto Western Hospital. To Dr. Alison Fleming, Dr. Robert

Inman, Dr. Zahi Touma, and Laura Passalent, thank you for your humility, generosity, mentorship, and for taking me in and giving me the foundation, inspiration, and opportunities to successfully pursue a career as a research scientist. I continue to look up to you and the example of excellence that you set as scientists.

Thank you to my classmates and the incredible staff at the Department of Health Research Methods, Evidence, and Impact, you have helped to make this experience that much more joyful. The end is bittersweet, and I feel extremely privileged to be an alumna of this historic and innovative department. I will carry the torch forward.

To my dearest friends, thank you for being my cheerleaders always, and through all of the moments that have led up to this point. Thank you for the check-ins and pep talks, the late nights and study sessions, the much needed breaks, and the adventures. Your encouragement and optimism have fueled me along the way.

I could not have accomplished this work without my family, who have consistently grounded and uplifted me, through the many ups and downs and everything in between. Thank you, Simon, Eileen, Dave Sr. and Tyler. Above all, I am indebted to my mother, Olena, for your incredible resilience, love, and selflessness, and the countless sacrifices you have made to get me here, сердечно дякую мама.

To my partner, Dave, I am forever grateful for your incalculable and unwavering patience, and for entertaining this segment of our lives with good humour. Thank you for everything, for your love, and for believing in me, I dedicate this work to you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CDSR:	Cochrane Database for Systematic Reviews
CENTRAL:	Cochrane Central Register of Controlled Trials
CI:	confidence interval
CINAHL:	Cumulative Index to Nursing and Allied Health Literature
CONSORT:	Consolidated Standards of Reporting Trials
EBR:	Evidence-Based Research
EQUATOR:	Enhancing the QUAlity and Transparency Of health Research
IQR:	interquartile range
KT:	knowledge translation
MeSH:	Medical Subject Headings
MISTIC:	MethodologIcal STudy reportIng Checklist
MR:	methodological review
PAFS:	pilot and feasibility studies
PRISMA:	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P:	Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
RCT:	randomized controlled trial
SD:	standard deviation
STROBE:	Strengthening the Reporting of Observational Studies in Epidemiology
SWAT/SWAR:	Studies Within a Trial/Studies Within a Review

Declaration of Academic Achievement

This dissertation is structured as a "sandwich thesis" which combines three independent studies with four manuscripts that are either published or prepared for submission in peer-reviewed journals. The results of the first two studies (Chapters 2 and 3) and the protocol (Chapter 4) outlining the third study have been published. The following are contributions of Daeria Lawson, the primary author, to all of the papers included in this thesis: study conception and development of the research questions; study design including writing the protocols and statistical analysis plans; data extraction and management; conduct of the statistical and qualitative analyses; interpretation of findings; writing all of the manuscripts including figures, tables, and supplementary materials; revision of manuscripts based on co-author feedback; and submission of manuscripts including responding to reviewers' comments. The contributions of co-authors vary for each study, detailed within each chapter, and include but are not limited to review of the study design; acquisition, management, and interpretation of data; and preparation and critical review of the manuscripts. The work in this thesis was conducted between Winter 2018 and Summer 2023.

CHAPTER 1: Introduction

1.1 Introduction

Health research addresses questions about the aetiology, progression, and prevention of disease, and the effects of treatments for diseases. However, and despite best efforts, errors in health research occur for a multitude of reasons (Brown, Kaiser, & Allison, 2018; Gluud & Hilden, 2009; Naylor, 1997). Arising from health research, methodological research aims to investigate strengths and limitations in the practice of health research. This is ultimately done to ensure or improve the way that we conduct health research such that the evidence is sound, and so that investigators can preserve the original intent of their research (Haynes, 2006). This includes improving the lives of patients while minimizing the physical, social, mental, and other harms as a result of participating in research. The following sections outline the history and purpose of methodological studies, issues in the field, and the layout of the thesis.

Brief history of methodological research

Many different study designs are referred to as methodological studies, and this is highly contextual and varies by scientific field. The scope of this thesis is methodological studies in health research with humans (biomedical or clinical), henceforth referred to as 'methodological studies' for brevity. We define these studies as those that synthesize evidence about research study practices. For the purpose of this thesis, the working definition of research study practices (or research practice) encompasses the design, conduct, analysis, and reporting or dissemination of health research.

In a recent study, Puljak et al. reported that the first instance of the term 'meta-epidemiology' (i.e., a type of, or synonym for methodological studies) appears in a 1997 editorial by David Naylor (Naylor, 1997; Puljak, Makaric, Buljan, & Pieper, 2020). This usage referred specifically to studies employing meta-regression techniques to investigate effect sizes as a result of trial design (Sterne et al., 2002). Dechartres et al. have stated that the first methodological study was published in 1929 by Dunn (Dunn, 1929) to assess statistical logic in a group of physiological articles (Dechartres, Charles, Hopewell, Ravaud, & Altman, 2011). Overall, it is clear that this scientific practice has existed for many decades, if perhaps not explicitly called out by name until recently.

Previous reviews have well delineated the types of questions that methodological studies can address including topics on performing, communicating, verifying, and rewarding research (Ioannidis, Fanelli, Dunne, & Goodman, 2015). These studies have been instrumental in inciting important developments in the design, conduct, analysis, and reporting of health research. This includes the conduct, evaluation, and reporting of randomized trials, observational studies, and systematic reviews (Altman & Simera, 2016; Buscemi, Hartling, Vandermeer, Tjosvold, & Klassen, 2006; Guyatt et al., 2008; Ioannidis et al., 2014; Moher, Jones, Lepage, & Group, 2001; Moher, Liberati, Tetzlaff, Altman, & Group, 2009; Schulz, Altman, Moher, & CONSORT Group, 2010; Schulz, Chalmers, Hayes, & Altman, 1995; von Elm et al., 2008; Wolff et al., 2019) to name a few, as well as access to research itself (Chan et al., 2014).

Undeniably these methodological efforts are necessary; they have promoted rigour and an evidence-based approach to health research practice (Chalmers & Glasziou, 2009;

Ioannidis et al., 2015). Rigour in health research is necessary since possession of scientific knowledge should be based not just on beliefs but reliable methods; and in every case, researchers should be able to justify their methods and knowledge claims (de Ridder, 2019; Glasziou et al., 2014). In the context of health research, human lives depend on efficient and reliable evidence (Ioannidis et al., 2015). Rigour can be achieved by involving patients in research, designing studies by referring to systematic reviews, controlling for biases in studies, and describing interventions adequately (Chalmers & Glasziou, 2009; Lund et al., 2016; Tugwell & Knottnerus, 2018). Rigour helps to reduce "research wastefulness" which is essentially the undertaking of research that is hampered by poor methodology. Methodological shortcuts, compromises, and omissions harm the institution of health science. The consequences are studies that are non-replicable, biased, or unpublished (Glasziou et al., 2014) among others. These and other issues have been reported over the years, and leading experts and global signatories such as the Open Science Framework, the San Francisco Declaration on Research Assessment (Center for Open Science, 2012; The American Society for Cell Biology, 2013), and others have made recommendations for researchers, institutions, and funders on ways to improve. However, experts suggest that there is still work to be done (von Niederhausern, Guyatt, Briel, & Pauli-Magnus, 2018), stating that "research waste is still a scandal" as recently as 2018 (Glasziou & Chalmers, 2018).

Issues with methodological research

The production of methods guidance is a common undertaking by research methodologists (Hirt et al., 2022). To wit, many guides have been published for

conducting and interpreting methodological studies in health research (Clarke, 2020; Hennessy, Johnson, & Keenan, 2019; Khalil & Munn, 2023; Lilford et al., 2001; Luchini et al., 2021; Martin et al., 2020; Mbuagbaw, Lawson, Puljak, Allison, & Thabane, 2020; Moustgaard et al., 2020; Munn, Stern, Aromataris, Lockwood, & Jordan, 2018; Sterne et al., 2002; Sutton, Clowes, Preston, & Booth, 2019; Zhang, 2016). This indicates interest among methodologists and the broader research community to perform such studies. Since these guides have the potential to influence the conduct of methodological studies, similarly, they also influence naming conventions. Collectively, these guides have referred to methodological studies as 'meta-epidemiological review/study', 'meta-research', 'methodological review/study', 'research-on-research', and other variations. These terms have all been used to describe a similar concept and study design. This ambiguity has been highlighted as early as in 1988 regarding the confusion of meta-analyses with 'methodology reviews' (Cunningham, 1988), and further established in a recent review (Puljak et al., 2020). This also suggests that there is no consensus, and to date, no prior attempts have been made to reach consensus on terminology by experts.

Not including studies published before the existence of any methods guidance, the above cited guidance papers inadvertently promote a variety of reporting styles for an analogous type of study. This results in a downstream effect which muddies the waters and complicates the search and retrieval process for these studies. To the best of our knowledge, with the exception of some topic specific databases (e.g., Cochrane Methodology Reviews, Core Outcome Measures in Effectiveness Trials), no comprehensive database exists for methodological studies. Investigator-led and project-

specific databases have been generated, but access is generally limited to study team members and collaborators. In some cases, access is considered upon request, and usually after study results have been published. For example, a group of researchers have produced a database of several, related meta-epidemiological studies for the purpose of pooling data for future research (Savovic et al., 2010). Currently, the lack of a dedicated database, and limited Medical Subject Headings (MeSH) in existing databases, complicates the identification of methodological studies.

Based on a recent study, authors found that search strategies based on titles, abstracts, and keywords were not sensitive enough to identify methodological studies (Penning de Vries, van Smeden, Rosendaal, & Groenwold, 2020). Others have found it challenging to distinguish the primary reports included in methodological studies (i.e., randomized trials) from the "meta-research" studies themselves (Nicholls, McDonald, McKenzie, Carroll, & Taljaard, 2022), as well as the finding that methodological studies of reporting were poorly reported (Santo et al., 2023). Based on several Cochrane Methodology Reviews, authors reported that it was difficult to find methodological studies and methods trials (Brueton et al., 2013; Odgaard-Jensen et al., 2011; Preston et al., 2016; Turner et al., 2012). Authors cited unclear reporting, lack of suitable search terms and inefficient search strategies, and a need to make search results manageable due to the irrelevance resulting from inefficient search strategies. Just as in systematic reviews, when searching for methods studies that evaluate pilot studies for example, one needs to specify the concept for the subject (i.e., pilot studies) in addition to terms for the concept of methodology. This can result in a large volume of records for screening. It is

expected that this aspect of screening and searching can be better automated with the application of machine learning methods (once a taxonomy is established), but ultimately researchers are still faced with final decision-making in such procedures.

These and other examples illustrate some of the difficulties in accessing methodological research, even by experts in literature searching and review methodology. If experts struggle with accessing methodological studies, how can less experienced end-users such as students, for example, do so? Albert Camus, the French-Algerian philosopher and Nobel laureate, was correct in stating, "To name things wrongly is to add to the misfortune of this world" (Camus, 1944).

Outline of thesis and issues addressed

The primary research question for this thesis is: how are methodological studies labelled and reported in the health research literature? With the hopes of answering this multi-faceted question, the following is an overview of the studies in this thesis which comprise a "sandwich" of four papers (Chapters 2 to 5). Throughout this thesis, methodological research of pilot and feasibility studies (PAFS) is referred to as a case study or specific area of focus. This is done with the intent to illustrate the previously mentioned challenges with methodological studies in a more digestible manner.

Chapter 2

Study design or article type: methods guidance

Background to case study of a methodological issue:

Pilot and feasibility studies help us to evaluate research before we commit to full

studies assessing the effectiveness, efficacy, and safety of treatments on a large number of patients. As a result, experts have recommended that these studies not be used to provide information on treatment efficacy (Thabane et al., 2010). However, this practice and other issues such as lack of a priori feasibility outcomes and progression criteria, and inaccurate naming conventions to disguise small studies as pilots persist (Arain, Campbell, Cooper, & Lancaster, 2010; Mbuagbaw et al., 2019; Shanyinde, Pickering, & Weatherall, 2011).

Objective and summary:

The primary objective of this study was to educate physiatrists on the appropriate methodology for the design, conduct, analysis, and reporting of PAFS. First, current practices were evaluated by way of a methodological study of the literature to examine researchers' objectives, and how and why they carried out PAFS. After key issues (e.g., missing specification of feasibility outcomes) had been identified, good practice recommendations were provided, tailored to rehabilitation research. This work also illustrates the types of issues that methodological studies can uncover, which is described in the first part of the study, before the methods guidance section.

Chapter 3

Study design or article type: pilot study of a review

Background:

As previously described, methodological studies that are otherwise similar in their intent (i.e., to synthesize evidence about research practice) are characterized by (1) inconsistent nomenclature, and (2) inconsistent reporting styles.

Objective and summary:

This study evaluated how to capture methodological studies (referred to as 'methodological reviews' in this chapter) in literature databases. This involved the development and pilot testing of a targeted search strategy with identified terms to yield methodological studies. The sensitivity and specificity of the search strategy was calculated, and progression criteria to a full review were prespecified. In addition, various methodology features were summarized from the sample of retrieved methodological studies. This work indicated the need for a concerted effort to streamline reporting and naming conventions.

Chapter 4

Study design or article type: protocol for reporting guideline development

Background:

Based on the findings of the pilot study (Chapter 3) and previous research, we registered the development of the MethodologIcal Study reportIng Checklist (MISTIC), a new reporting guideline for methodological studies that synthesize evidence about health research practices, i.e., design, conduct, analysis, and reporting. As part of this exercise, a review of the literature would allow us to identify previous reporting guidelines, if any. A review of the international authority on reporting guidelines—Enhancing the QUAlity and Transparency Of health Research (EQUATOR)—confirmed that no formal reporting guideline has been previously registered (EQUATOR Network, 2019).

Objective and summary:

This article outlines the procedures to (1) define and harmonise the names

describing methodological studies, and (2) develop a reporting guideline for methodological studies in human health research. The impact of methodological studies on health research, implications for research practice from this project, and approaches to knowledge translation are discussed. Stakeholders from Canada, Croatia, Denmark, Germany, South Africa, Switzerland, United Kingdom, and United States with expertise in health research methodology, clinical epidemiology, journalology, statistics, and reporting guideline development were recruited. These individuals were targeted because they are representative of the researchers conducting and publishing methodological studies. Involving key stakeholders in guideline development helps to support its relevance, wider acceptance, and future uptake (Esmail, Moore, & Rein, 2015). This protocol is structured in accordance with the best practices for the development of reporting guidelines (Moher, Schulz, Simera, & Altman, 2010).

Chapter 5

Study design or article type: review/methodological study

Background:

As part of the first phase of development of the MISTIC reporting guideline, we conducted a review of methodological studies, in line with the previously mentioned working definition. This study was designed on the basis of the findings from the pilot study (Chapter 3).

Objective and summary:

This study aimed to characterize methodological studies based on (1) the names used to describe them, (2) the reporting practices of authors of methodological studies

including any cited reporting guidance, and (3) their methodology. As part of the full review described in the protocol (Chapter 4), this paper summarizes methodological studies that specifically evaluated PAFS (i.e., same as those evaluated in Chapter 2). The third listed study aim—the investigation of methodology within methodological studies—supports the development of a reporting guideline by allowing us to stratify and develop a taxonomy of methodological studies. Accordingly, this can help to gauge the appropriateness of existing or a new reporting guideline on the basis of the methodological studies' study designs (i.e., should they vary sufficiently to warrant distinct reporting guidance).

1.2 References

Altman, D. G., & Simera, I. (2016). A history of the evolution of guidelines for reporting medical research: the long road to the EQUATOR Network. *J R Soc Med, 109*(2), 67-77. doi:10.1177/0141076815625599

The American Society for Cell Biology. (2013). San Francisco Declaration on Research Assessment. Retrieved from https://sfdora.org/read/

Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Med Res Methodol, 10*, 67. doi:10.1186/1471-2288-10-67

Brown, A. W., Kaiser, K. A., & Allison, D. B. (2018). Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proc Natl Acad Sci U S A, 115*(11), 2563-2570. doi:10.1073/pnas.1708279115

Brueton, V. C., Tierney, J., Stenning, S., Harding, S., Meredith, S., Nazareth, I., & Rait, G. (2013). Strategies to improve retention in randomised trials. *Cochrane Database Syst Rev*(12), MR000032. doi:10.1002/14651858.MR000032.pub2

Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., & Klassen, T. P. (2006). Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol, 59*(7), 697-703. doi:10.1016/j.jclinepi.2005.11.010

Camus, A. (1944). Sur une philosophie de l'expression. In P. Seghers (Ed.), *Poésie* (pp. 44). Paris.

Center for Open Science. (2012). Open Science Framework (OSF). Retrieved from https://osf.io

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet, 374*(9683), 86-89. doi:10.1016/S0140-6736(09)60329-9

Chan, A. W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gotzsche, P. C., . . . van der Worp, H. B. (2014). Increasing value and reducing waste: addressing inaccessible research. *Lancet, 383*(9913), 257-266. doi:10.1016/S0140-6736(13)62296-5

Clarke, M. (2020). Guide to the contents of a Cochrane Methodology protocol and review. In H. J. P. & S. Green (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.1.0 ed.): The Cochrane Collaboration.

Cunningham, A. S. (1988). Meta-analysis and methodology review: what's in a name? *J Pediatr, 113*(2), 328-329. doi:10.1016/s0022-3476(88)80274-9

de Ridder, J. (2019). *How many scientists does it take to have knowledge?* (1st ed.). New York, NY: Routledge.

Dechartres, A., Charles, P., Hopewell, S., Ravaud, P., & Altman, D. G. (2011). Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol, 64*(2), 136-144. doi:10.1016/j.jclinepi.2010.04.015

Dunn, H. L. (1929). Application of statistical methods in physiology. *Physical Review, 9*, 275-398.

EQUATOR Network. (2019). The EQUATOR Network register of reporting guidelines under development. Retrieved from https://www.equator-network.org/library/reporting-guidelines-under-development/

Esmail, L., Moore, E., & Rein, A. (2015). Evaluating patient and stakeholder engagement in research: moving from theory to practice. *J Comp Eff Res, 4*(2), 133-145. doi:10.2217/cer.14.79

Glasziou, P., Altman, D. G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., . . . Wager, E. (2014). Reducing waste from incomplete or unusable reports of biomedical research. *Lancet, 383*(9913), 267-276. doi:10.1016/S0140-6736(13)62228-X

Glasziou, P., & Chalmers, I. (2018). Research waste is still a scandal—an essay by Paul Glasziou and Iain Chalmers. *BMJ, 363*, 1-3. doi:10.1136/bmj.k4645

Gluud, C., & Hilden, J. (2009). Povl Heiberg's 1897 methodological study on the statistical method as an aid in therapeutic trials. *Prev Med, 48*(6), 600-603. doi:10.1016/j.ypmed.2009.04.001

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., . . . GRADE Working Group. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336*(7650), 924-926. doi:10.1136/bmj.39489.470347.AD

Haynes, R. B. (2006). *Clinical epidemiology: how to do clinical practice research* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.

Hennessy, E. A., Johnson, B. T., & Keenan, C. (2019). Best Practice Guidelines and Essential Methodological Steps to Conduct Rigorous and Systematic Meta-

Reviews. *Appl Psychol Health Well Being, 11*(3), 353-381. doi:10.1111/aphw.12169

Hirt, J., Ewald, H., Lawson, D. O., Hemkens, L. G., Briel, M., & Schandelmaier, S. (2022). A systematic survey of methods guidance suggests areas for improvement regarding access, development, and transparency. *J Clin Epidemiol, 149*, 217-226. doi:10.1016/j.jclinepi.2022.05.005

Ioannidis, J. P., Fanelli, D., Dunne, D. D., & Goodman, S. N. (2015). Meta-research: Evaluation and Improvement of Research Methods and Practices. *PLoS Biol, 13*(10), e1002264. doi:10.1371/journal.pbio.1002264

Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., . . . Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *Lancet, 383*(9912), 166-175. doi:10.1016/S0140-6736(13)62227-8

Khalil, H., & Munn, Z. (2023). Guidance on conducting methodological studies – an overview. *Curr Opin Epidemiol Public Health, 2*, 1-5. doi:10.1097/PXH.0000000000000013

Lilford, R. J., Richardson, A., Stevens, A., Fitzpatrick, R., Edwards, S., Rock, F., & Hutton, J. L. (2001). Issues in methodological research: perspectives from researchers and commissioners. *Health Technol Assess, 5*(8), 1-57. doi:10.3310/hta5080

Luchini, C., Veronese, N., Nottegar, A., Shin, J. I., Gentile, G., Granziol, U., . . . Solmi, M. (2021). Assessing the quality of studies in meta-research: Review/guidelines

on the most important quality assessment tools. *Pharm Stat, 20*(1), 185-195. doi:10.1002/pst.2068

Lund, H., Brunnhuber, K., Juhl, C., Robinson, K., Leenaars, M., Dorch, B. F., . . . Chalmers, I. (2016). Towards evidence based research. *BMJ, 355*, i5440. doi:10.1136/bmj.i5440

Martin, G. P., Jenkins, D. A., Bull, L., Sisk, R., Lin, L., Hulme, W., . . . Predictive Healthcare Analytics, G. (2020). Toward a framework for the design, implementation, and reporting of methodology scoping reviews. *J Clin Epidemiol, 127*, 191-197. doi:10.1016/j.jclinepi.2020.07.014

Mbuagbaw, L., Kosa, S. D., Lawson, D. O., Stalteri, R., Olaiya, O. R., Alotaibi, A., & Thabane, L. (2019). The reporting of progression criteria in protocols of pilot trials designed to assess the feasibility of main trials is insufficient: A meta-epidemiological study. *Pilot and Feasibility Studies, 5*(1). doi:doi:

Mbuagbaw, L., Lawson, D. O., Puljak, L., Allison, D. B., & Thabane, L. (2020). A tutorial on methodological studies: the what, when, how and why. *BMC Med Res Methodol, 20*(1), 226. doi:10.1186/s12874-020-01107-7

Moher, D., Jones, A., Lepage, L., & Group, C. (2001). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA, 285*(15), 1992-1995. doi:10.1001/jama.285.15.1992

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol, 62*(10), 1006-1012. doi:10.1016/j.jclinepi.2009.06.005

Moher, D., Schulz, K. F., Simera, I., & Altman, D. G. (2010). Guidance for developers of health research reporting guidelines. *PLoS Med, 7*(2), e1000217. doi:10.1371/journal.pmed.1000217

Moustgaard, H., Jones, H. E., Savovic, J., Clayton, G. L., Sterne, J. A., Higgins, J. P., & Hrobjartsson, A. (2020). Ten questions to consider when interpreting results of a meta-epidemiological study-the MetaBLIND study as a case. *Res Synth Methods, 11*(2), 260-274. doi:10.1002/jrsm.1392

Munn, Z., Stern, C., Aromataris, E., Lockwood, C., & Jordan, Z. (2018). What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Med Res Methodol, 18*(1), 5. doi:10.1186/s12874-017-0468-4

Naylor, C. D. (1997). Meta-analysis and the meta-epidemiology of clinical research. *BMJ, 315*(7109), 617-619. doi:10.1136/bmj.315.7109.617

Nicholls, S. G., McDonald, S., McKenzie, J. E., Carroll, K., & Taljaard, M. (2022). A review identified challenges distinguishing primary reports of randomized trials for meta-research: A proposal for improved reporting. *J Clin Epidemiol, 145*, 121-125. doi:10.1016/j.jclinepi.2022.01.013

Odgaard-Jensen, J., Vist, G. E., Timmer, A., Kunz, R., Akl, E. A., Schunemann, H., . . . Oxman, A. D. (2011). Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev, 2011*(4), MR000012. doi:10.1002/14651858.MR000012.pub3

Penning de Vries, B. B. L., van Smeden, M., Rosendaal, F. R., & Groenwold, R. H. H. (2020). Title, abstract, and keyword searching resulted in poor recovery of articles in systematic reviews of epidemiologic practice. *J Clin Epidemiol, 121*, 55-61. doi:10.1016/j.jclinepi.2020.01.009

Preston, N. J., Farquhar, M. C., Walshe, C. E., Stevinson, C., Ewing, G., Calman, L. A., . . . Todd, C. (2016). Strategies designed to help healthcare professionals to recruit participants to research studies. *Cochrane Database Syst Rev, 2*(2), MR000036. doi:10.1002/14651858.MR000036.pub2

Puljak, L., Makaric, Z. L., Buljan, I., & Pieper, D. (2020). What is a meta-epidemiological study? Analysis of published literature indicated heterogeneous study designs and definitions. *J Comp Eff Res, 9*(7), 497-508. doi:10.2217/cer-2019-0201

Santo, T. D., Rice, D. B., Amiri, L. S. N., Tasleem, A., Li, K., Boruff, J. T., . . . Thombs, B. D. (2023). Methods and results of studies on reporting guideline adherence are poorly reported: a meta-research study. *Journal of Clinical Epidemiology*. doi:https://doi.org/10.1016/j.jclinepi.2023.05.017

Savovic, J., Harris, R. J., Wood, L., Beynon, R., Altman, D., Als-Nielsen, B., . . . Sterne, J. A. (2010). Development of a combined database for meta-epidemiological research. *Res Synth Methods, 1*(3-4), 212-225. doi:10.1002/jrsm.18

Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Trials, 11*, 32. doi:10.1186/1745-6215-11-32

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA, 273*(5), 408-412. doi:10.1001/jama.273.5.408

Shanyinde, M., Pickering, R. M., & Weatherall, M. (2011). Questions asked and answered in pilot and feasibility randomized controlled trials. *BMC Med Res Methodol, 11*, 117. doi:10.1186/1471-2288-11-117

Sterne, J. A., Juni, P., Schulz, K. F., Altman, D. G., Bartlett, C., & Egger, M. (2002). Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med, 21*(11), 1513-1524. doi:10.1002/sim.1184

Sutton, A., Clowes, M., Preston, L., & Booth, A. (2019). Meeting the review family: exploring review types and associated information retrieval requirements. *Health Info Libr J, 36*(3), 202-222. doi:10.1111/hir.12276

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., . . . Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol, 10*, 1. doi:10.1186/1471-2288-10-1

Tugwell, P., & Knottnerus, J. A. (2018). Scientific integrity includes the rigour of the methods. *J Clin Epidemiol, 93*, v-vi. doi:10.1016/j.jclinepi.2017.12.023

Turner, L., Shamseer, L., Altman, D. G., Weeks, L., Peters, J., Kober, T., . . . Moher, D. (2012). Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in

medical journals. *Cochrane Database Syst Rev, 11*(11), MR000030. doi:10.1002/14651858.MR000030.pub2

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., & Initiative, S. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol, 61*(4), 344-349. doi:10.1016/j.jclinepi.2007.11.008

von Niederhausern, B., Guyatt, G. H., Briel, M., & Pauli-Magnus, C. (2018). Academic response to improving value and reducing waste: A comprehensive framework for INcreasing QUality In patient-oriented academic clinical REsearch (INQUIRE). *PLoS Med, 15*(6), e1002580. doi:10.1371/journal.pmed.1002580

Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., . . . Groupdagger, P. (2019). PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med, 170*(1), 51-58. doi:10.7326/M18-1376

Zhang, Z. (2016). Meta-epidemiological study: a step-by-step approach by using R. *J Evid Based Med, 9*(2), 91-97. doi:10.1111/jebm.12191

CHAPTER 2: A methodological study of pilot and feasibility studies and methods guidance

2.0 Preface

The following pages include the published study covering Chapter 2. Reprinted in accordance with copyright license and permission from Wolters Kluwer Health, Inc. This is a non-final version of an article published in final form in the American Journal of Physical Medicine & Rehabilitation. The policy permits the final peer-reviewed manuscript (i.e., accepted version) of the article to be reused in a thesis twelve months after the article's publication date.

Citation:

Lawson, D. O., Mellor, K., Eddy, S., Lee, C., Kim, K. H., Kim, K., . . . Thabane, L. (2022). Pilot and Feasibility Studies in Rehabilitation Research: A Review and Educational Primer for the Physiatrist Researcher. American Journal of Physical Medicine & Rehabilitation, 101, 372-383. doi:10.1097/PHM.0000000000001797

Pilot and feasibility studies in rehabilitation research: a review and educational primer for the physiatrist researcher

Daeria O. Lawson HBSc[1], Katie Mellor BSc MSc[2], Saskia Eddy BSc MSc[3], Christopher Lee BMSc MPH[4], Kwang Hyun Kim BHSc[4], Kunmo Kim BMSc[4], Lawrence Mbuagbaw MD MPH PhD[1,5], Lehana Thabane PhD[1,5]

[1] Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON Canada

[2] Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom

[3] Institute of Population Health Sciences, Queen Mary University of London, London, United Kingdom

[4] Department of Medicine, McMaster University, Hamilton, ON Canada

[5] Biostatistics Unit, Father Sean O'Sullivan Research Centre, St. Joseph's Healthcare Hamilton, Hamilton, ON Canada

Corresponding author: Lehana Thabane, Biostatistics Unit, Father Sean O'Sullivan Research Centre, 3rd Floor Martha Wing, Room H325, St. Joseph's Healthcare Hamilton, 50 Charlton Avenue East, Hamilton ON, L8N 4A6 Canada; Tel 1-905-522-1155 ext. 33720/34905; Fax 1-905-528-7386; E-mail thabanl@mcmaster.ca

Author disclosures: Lehana Thabane and Lawrence Mbuagbaw are on the editorial board of *Pilot and Feasibility Studies*. Lehana Thabane is one of the developers of the CONSORT extension for pilot and feasibility trials. The authors have not received any financial benefits for this work. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. This research has not been previously presented or published elsewhere in any form.

ABSTRACT

Pilot and feasibility studies are conducted early in the clinical research pathway to evaluate whether a future, definitive study can or should be done, and if so, how. Poor planning and reporting of pilot and feasibility studies can compromise subsequent research efforts. Inappropriate labelling of studies as pilots also compromises education. In this review, first a systematic survey of the current state of pilot and feasibility studies in rehabilitation research was performed, and second, recommendations were made for improvements to their design and reporting. In a random sample of 100 studies, half (49.5%) were randomized trials. Thirty (30.0%) and three (3.0%) studies used 'pilot' and 'feasibility' in the study title, respectively. Only one third (34.0%) of studies provided a primary objective related to feasibility. Most (92.0%) studies stated an intent for hypothesis testing. Although many (70.0%) studies mentioned outcomes related to feasibility in the methods, a third (30.0%) reported additional outcomes in the results and discussion only, or commented on feasibility anecdotally. The reporting of progression plans to a main study (21.0%) and progression criteria (4.0%) was infrequent. Based on these findings, it is recommended that researchers correctly label studies as a pilot or feasibility design based on accepted definitions, explicitly state feasibility objectives, outcomes and criteria for determining success of feasibility, justify the sample size, and appropriately interpret and report the implications of feasibility findings for the main future study.

KEY WORDS

feasibility, pilot studies, rehabilitation research, reporting

BACKGROUND

Pilot and feasibility studies in clinical research have been defined as studies that investigate the feasibility of methods and procedures to be used and evaluated in larger studies [1,2]. Due to the different definitions and interchangeable use of these terms in the literature, Eldridge et al. developed a consensus-based conceptual framework wherein:

- A *feasibility study* asks whether the future study can be done, should be done, and, if so, how; and
- A *pilot study* is a type of feasibility study that asks the same questions about feasibility as above, but the planned definitive study, or part of it, is conducted on a smaller scale [3].

As such, feasibility is the overarching concept [3]—a measure which can take on various forms depending on the feasibility objectives—and an essential component of pilot and feasibility studies. Therefore, "all pilot studies are feasibility studies but not all feasibility studies are pilot studies" [3].

Pilot and feasibility studies are valuable to ensure methodologically rigorous and higher quality future studies [4], and as a result, play a key role in helping to reduce research waste [5] when optimally planned. Thabane et al. classified the reasons for carrying out a pilot or feasibility study under four broad ventures [1]: 1) process – to evaluate the necessary steps of the main study, such as recruitment; 2) resources – to evaluate logistical aspects such budget, space, or time to complete questionnaires, all of which may impact the main study; 3) management – to evaluate both human- and data-related elements, such as the capacity of participating centres to manage study data; and

4) scientific – to determine appropriate dosing, treatment safety, or estimating the treatment effect, although there are divergent opinions about safety and efficacy estimates as reasons for performing a pilot study [6]. However, the emphasis is on preliminary "estimation" for the purposes of performing a sample size calculation for the main study, since the measurement of definitive efficacy is beyond the scope of pilot and feasibility studies.

In the context of physical medicine and rehabilitation, researchers have outlined a number of challenges unique to this field including: difficulties with recruitment in single settings; difficulties with blinding (participants and administrators) due to the nature of the interventions; participants' resistance to random assignment to interventions; difficulties with retaining participants in the control arm, and the unacceptability of withholding or delaying treatment (i.e., as a control); and the complexity of interventions and requirements for long-term follow-up, which further complicate multicenter initiatives [7,8]. Due to extensive issues with attrition in rehabilitation research—since many interventions require a great deal of time and effort from the patient—proposed future directions include efforts to better measure and enhance patient adherence and cooperation with data collection [7], which also happens to be a strength of, and best determined through pilot work. Additionally, owing to the complexity of rehabilitation research, pilot and feasibility studies are especially valuable since they focus on evaluating the failure of implementation and not the failure of the intervention [9].

In the past two decades, a number of methodological reviews have identified issues with the conduct and reporting of pilot and feasibility studies, citing

inconsistencies in naming conventions [4,10,11], a high emphasis on estimating and evaluating treatment effects [10,12], and lack of explicit plans or criteria for progression to a definitive study [10,12-14]. Although there has been a greater focus on the review of and guidance for randomized pilots informing future randomized trials, most published recommendations also apply to pilots with non-randomized designs. This paper has been structured from an educational perspective based on the following key aims: 1) to summarize current practices in the conduct and reporting of pilot and feasibility studies in rehabilitation research; 2) to identify the prevalence of the above key issues, and highlight knowledge gaps; and 3) to provide examples of best practices and recommendations for the design and reporting of pilot and feasibility studies specific to the field of rehabilitation.

METHODS

Study design

This is a review with a systematic survey and narrative synthesis of pilot and feasibility studies.

Search strategy

A search strategy was developed with the aid of a health sciences librarian. Medical subject headings were used for "pilot study" and "feasibility study" with a combination of keyword searching (see Supplemental Digital Content 1 for full search strategy). The search was carried out on July 31, 2020 and restricted to the top six journals in rehabilitation research, based on expert opinion, including the *American*

Journal of Physical Medicine and Rehabilitation, *American Journal of Sports Medicine*, *Archives of Physical Medicine and Rehabilitation*, *British Journal of Sports Medicine*, *European Journal of Physical and Rehabilitation Medicine*, and *PM&R Journal*. There were no restrictions by timeframe, language, or publication type.

Eligibility criteria

All inclusion and exclusion criteria are outlined in Table 1. Due to the interchangeable use of the terms "pilot" and "feasibility", studies were eligible if the terms were used differently from the current framework [3]. However, reviewers verified whether the study was not simply a small study and that some aspect of the report was suggestive of feasibility testing or piloting for a larger study.

Study selection

Three reviewers (DOL, KM, SE) independently screened the titles and abstracts of all retrieved articles to identify potentially relevant articles using *Rayyan* [15], a web-based systematic review platform. The same reviewers independently screened the full-texts of all relevant articles in *Covidence* [16], a web-based systematic review platform, to identify eligible studies for inclusion. Each abstract and full-text was reviewed twice (i.e., by two reviewers). All reviewers completed calibration exercises to pilot the screening forms at each stage (20 studies per reviewer).

Randomization

In order to evaluate a representative body of research, it was determined that a sample of 100 studies would be sufficient for an education-oriented review with no

planned statistical comparisons [17]. If fewer than 100 eligible studies were identified, reviewers planned to extract data from all of them. Software-generated random numbers—using the Mersenne Twister algorithm—were automatically assigned to all eligible studies in *Microsoft Excel* [18]. Studies were sorted by assigned number and studies with random numbers 1-100 were selected for data extraction. Similarly, the order of data extraction was randomized for the sample of 100 studies. One reviewer (DOL) carried out the above randomization procedures.

Data extraction

Six reviewers (DOL, KM, SE, CL, KHK, KK) independently extracted data from the included studies into standardized forms in *Microsoft Excel* [18]. Each study was evaluated twice (i.e., by two reviewers). To minimize information bias (i.e., differential misclassification), reviewers captured all relevant study design information as reported in the appropriate sections (e.g., the outcomes as listed in the methods; aims to evaluate feasibility and intent for a larger study as listed in the objectives or methods). As such, the data collection forms were organized by location in the report from where information was drawn (e.g., objectives, methods, results, or discussion sections). This structure not only reflected universally minimal reporting standards for key study design details, but also aligns with clearly planned objectives by authors. Data was extracted for the following: bibliometrics (e.g., author information, country), reported description of study design (pilot, feasibility), study design characteristics (e.g., registration, funding, condition investigated, intervention and comparator type, sample size, plans to estimate treatment effect), and general items from the Consolidated Standards of Reporting Trials

(CONSORT) extension for randomized pilot and feasibility trials [19] that are not specific to a randomized design (e.g., if the objective was related to feasibility, types of feasibility outcomes investigated, conclusions with regard to feasibility of a main study). Reviewers did not check cited protocols, trial registrations or supplemental files for additional information. All reviewers completed calibration exercises to pilot the data extraction forms (minimum of five studies per reviewer).

Due to the difficulties reviewers faced in identifying explicit statements for the two data fields relating to whether a) study authors discussed the implications of their study's feasibility findings, and b) conclusions with regard to feasibility, the original criteria for classifying studies (1-stop, definitive study not feasible; 2-continue but modify protocol; 3-continue without modifications but close monitoring required; 4-continue without modifications; 5-not reported; 6-other) were operationalized into traffic light indicator categories during data analysis (1-stop, not feasible; 2-caution, feasible with changes; 3-start, feasible as is; 4-not reported). This field incorporated verbatim quotations regarding study design suggestions and planned undertakings, and author's intentions to proceed to or inform future studies had to be clear for 1-3 ratings. Any studies without explicit suggestions to alter the current study design for larger studies were classified 'feasible as is'. Generic statements about the need for future research, or comments addressing common study design limitations not specific to pilot and feasibility studies (e.g., the resulting small sample size, and therefore need for larger studies, or lack of a control group and therefore randomized trials are needed) were not considered appropriate conclusions regarding the feasibility of or progression to future

research. Such conclusions were classified as 'not reported' unless authors specifically linked one of these factors to feasibility (e.g., limited by small sample size for efficacy conclusions and data will be used to estimate the sample size for a future study).

Two reviewers (DOL and KM) also extracted editorial policy information from each journal's "For Authors" (or equivalent) section. Reviewers classified journals based on the degree of study design reporting guidance available: high (i.e., require authors to append appropriate reporting checklist), moderate (i.e., require authors to refer to reporting guidelines broadly or as needed), low (i.e., general endorsement of good reporting practices without specific guidance by study design), and none (i.e., no specific guidance aside from article type submission guidelines).

All conflicts during screening and data extraction were resolved through discussion between each pair of reviewers, and if no consensus could be reached, a third reviewer (LM or LT) adjudicated.

Data analysis

Data was summarized using explanatory text in tables including verbatim quotations for illustrative purposes. Descriptive statistics using counts (percentage) for categorical data and means (standard deviation [SD]) for continuous data with 95% confidence intervals [CI] were computed in *IBM SPSS Statistics v. 24.0 for Macintosh* [20].

RESULTS

Of the 252 eligible studies, 100 randomly selected studies with 3106 patients were included and summarized below (see Supplemental Digital Content 2 for list of included

studies). The flow of articles is illustrated in Figure 1.

General characteristics

General characteristics of the included studies are outlined in Table 2. The majority (52.0%) of studies were published in *Archives of Physical Medicine and Rehabilitation*, and in the most recent two decades with 87 (87.0%) published between 2005-2020. A majority (53.5%) of studies were from North America, followed by Europe (26.7%). Most studies used quantitative methodology with 50 (49.5%) randomized trials and 41 (40.6%) uncontrolled before-after studies. The most common condition investigated was stroke (20.4%) and the number of patient participants analyzed ranged from 3 to 194. In addition to patient participants, studies also evaluated clinicians (3.0%), study personnel (1.0%) and caregivers (1.0%). Movement or exercise-based interventions (48.0%) were the most prevalent.

Feasibility-related characteristics

All feasibility-related characteristics of the included studies are outlined in Table 3 and summarized below.

1. Feasibility objectives:

About a third (34.0%) of studies stated a primary objective related to feasibility (e.g., the goal of the study is 'to examine (1) feasibility...', 'to establish if the treatment could be safely tolerated', 'to evaluate patient acceptance and satisfaction with the treatment...'). In a post hoc comparison, studies reporting a primary feasibility objective had lower proportions of a randomized study design (38.2% and 56.1%), reporting a

sample size estimation (17.6% and 30.3%), and reporting additional feasibility outcomes in the results or discussion (17.6% and 36.4%) compared to studies that did not report a primary feasibility objective, respectively. In contrast, studies reporting a primary feasibility objective had higher proportions of mentioning that the current study was conducted in preparation for a main or larger study (41.2% and 9.1%), mentioning progression plans with(out) criteria (41.2% and 3.0%), and providing a discussion of the implications of feasibility findings (91.2% and 39.4%) compared to studies that did not report a primary feasibility objective, respectively. The comparison is illustrated in Figure 2.

2. Feasibility outcomes:

The majority (70.0%) of studies reported outcomes related to feasibility a priori in the methods, with patient adherence (19.5%) and safety (17.6%) being most common. Almost a third (30.0%) of studies reported results for additional outcomes related to feasibility in the results or discussion sections with the most common being patient adherence (10.4%) and safety (7.8%). Additionally, the following outcomes related to feasibility were reported or commented on anecdotally in four (4.0%) studies: patient acceptance (2), patient feedback (2), and clinician feedback (1) on the intervention.

3. Sample size estimation:

A sample size estimation was reported in a quarter (25.0%) of studies. Among these, only one (0.9%) study justified the calculation based on feasibility outcomes. Most studies justified their calculations based on the intervention effect (15.8%) or similar

studies (16.7%). Two additional studies reported post hoc calculations, one of which stated that a 'power analysis was performed' although no calculation or values were reported.

4. Progression plans to a main study:

Authors mentioned that the current study was done to inform the design of a main or larger study in about one fifth (21.0%) of studies. Among these, criteria to measure the success of at least one feasibility outcome in order to progress were provided in three (3.0%) studies. Authors of an additional study provided criteria to measure success but from the perspective of implementation rather than plans for future research (noting that the pilot was done in preparation for a larger study only in the discussion). Examples of progression criteria and associated results as reported by authors are summarized in Table 4.

5. Estimation of treatment effect:

Authors of most (92.0%) studies stated an intent for hypothesis testing or to provide an estimate of the treatment effect. Three quarter (75.0%) of studies that did not state any plans to estimate the treatment effect had feasibility as their primary objective (not shown). Authors of one study stated an intent to measure the treatment effect to be reported in a separate publication.

6. Conclusions about feasibility:

The implications of feasibility findings on the main or larger study were addressed in over half (57.0%) of the studies. No authors concluded that their study was not

feasible. Based on the reviewers' criteria, almost half (43.0%) of studies were classified as feasible with modifications, and less than a fifth (14.0%) were classified feasible as is. Examples of conclusions related to feasibility as reported by authors are summarized in Supplemental Table 1 (Supplemental Digital Content 3).

7. Reporting guideline use:

References to reporting guidelines were rare with only nine (9.0%) studies citing a specific guideline. Among these, the CONSORT statement was most frequently cited (5.0%), followed by the CONSORT pilot and feasibility extension (2.0%). Authors of non-randomized studies cited the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline (1.0%), and the Template for Intervention Description and Replication (TIDieR) checklist (1.0%).

Journal editorial policies

Among the six journals, four (66.7%) required authors to append an appropriate reporting checklist (e.g., CONSORT) during submission (high guidance), one (16.7%) required authors to refer to reporting guidelines as needed (moderate guidance), and one (16.7%) did not provide any reporting guidance (i.e., aside from 'article type' submission guidelines). Among the high-moderate level guidance journals, one (16.7%) provided links for downloadable checklist forms, three (50.0%) provided links to the specific guideline or Enhancing the QUAlity and Transparency Of health Research (EQUATOR) websites, and one (16.7%) provided no links or downloadable forms.

DISCUSSION

In this review of rehabilitation pilot and feasibility studies, only a third (34.0%) of studies reported a primary feasibility objective. A third (30.0%) did not mention any preplanned feasibility outcomes, and a third (30.0%) reported additional feasibility outcomes elsewhere in the manuscript for the first time (or commented on feasibility anecdotally). Similarly, the reporting of progression plans to a definitive study was infrequent (21.0%). In the majority of included studies, authors discussed the need for future randomized trials to further evaluate the intervention and highlighted general limitations such as: small sample size, or lack of a control group (in non-randomized, single-arm designs) which contributed to uncertainties in the attribution of the measured effect to the intervention. This is supplemented by the finding that 92% of all studies planned to estimate the treatment effect, with few (9.0%) randomized studies performing an intention-to-treat analysis. Pilot and feasibility studies are not intended to assess efficacy or effectiveness, and incorrect labelling of small, underpowered studies as pilots increases the risk of spurious findings and misinterpretations of efficacy, which can impact future evidence synthesis efforts if these studies are included in meta-analyses as small studies rather than pilots.

The above findings illustrate that confusion persists in the field with regard to study design. Poor planning of studies also undermines the significant patient investments in clinical research. A requirement for reporting details such as the intent for a definitive study in the objectives or methods sections of a report may seem pedantic to some; however, this is important to be able to delineate whether the reported feasibility

assessment of a future study was preplanned by the authors. Readers should always be told explicitly with no room for inferring what was done [21] and why. For example, in one study authors suggested that an outcome (pain intensity) would be used to assess feasibility, however, it was unclear how this outcome was used to measure feasibility up until the discussion (i.e., tolerance), since this can also be used as a clinical measure of treatment effect [22].

Incorrect labelling of studies and other reporting issues also hamper education efforts. It is important that researchers carefully consider and incorporate feasibility objectives, outcomes and progression criteria when planning pilot and feasibility studies. If researchers set out to measure feasibility without further comments about its implications, this would be a flag. Similarly, certain attributes including only general discussions of study design limitations, discussions mostly concerned with efficacy, or first mentions of feasibility objectives for the current study in the discussion would also suggest that the study was not initially preplanned as a pilot or feasibility study (i.e., and labelled as such after an unexpectedly small sample was obtained). A more concerning scenario is one in which the label of 'pilot' or 'feasibility' may have been added on at the manuscript submission or editorial stage. Examples of inadequate feasibility conclusions including probable instances of small studies labelled as pilots are outlined in Supplemental Table 1 (Supplemental Digital Content 3).

It is not surprising that there was an equal number of randomized and non-randomized studies in this sample of rehabilitation research, especially since complex interventions are common. However, in both approaches to design, there is room for

improvement in adequately planning for and incorporating objectives and measures of feasibility, as well as explicitly addressing plans for a future definitive study. Although the majority of developments in pilot and feasibility study methodology and reporting have targeted studies preceding randomized trials, most recommendations apply to non-randomized designs. The following section highlights key issues identified with the design and reporting of external pilot and feasibility studies in rehabilitation research, and presents resources to assist physiatrists from planning to reporting pilot and feasibility studies. Readers are referred to a review by Avery et al. for additional considerations for internal pilot trials [23].

Key issues and recommendations

1. General resources for designing pilot and feasibility studies

There are a number of resources for researchers who are interested in designing pilot or feasibility studies. The CONSORT pilot and feasibility trials working group developed a website with guidance for trialists, including information about design, analysis and reporting [24]. Similarly, Trial Forge is a resource which focuses on reviewing methodological topics and summarizing methodological evidence to help researchers design more efficient trials [25]. "Evidence Packs" can be downloaded, which include a curated set of tips, guidance, and relevant articles about various trial design topics including pilot and feasibility studies (although at the time of this writing this section was being updated). Previously cited research [1,3,10] outlines minimum considerations when designing a randomized pilot or feasibility study. Additional considerations for non-randomized designs have also been published [26]. The use of uncontrolled before-after

studies as an adequate source of evidence for causation has been discouraged by some groups [27], while others have stated that randomized trials pose limitations to addressing certain types of clinical questions unique to rehabilitation [28]. Discouragement is largely due to challenges with confounding and uncertainties in attribution of the effect, conclusions which have also been observed in studies in this review. Authors need to carefully consider and justify all decisions for a chosen study design, while taking into account the limitations of pilot and feasibility studies for establishing efficacy and causation.

2. Naming conventions

At the full-text screening stage, it was apparent that many of the studies that were classified as "wrong study design" by reviewers were in fact small trials labelled as pilot studies. Consequently, there are various justifications for labelling a study as a pilot or feasibility study, especially based on the persisting misconception that a small sample size defines a pilot study. Such inappropriate reasoning or apparent post hoc decisions (e.g., "funding ended and so the study was considered a pilot", "small sample size limited the current study to a feasibility study", or "due to the small number of patients that participated, we regarded this as a pilot") were not uncommon in this review and during screening. To outline the subtle differences, based on the proposed framework by Eldridge et al. [3], pilot studies fall under the umbrella of feasibility studies, and ask the same questions about feasibility (whether the future study can be done, should be done, and, if so, how) but the planned definitive study, or part of it, is conducted on a smaller scale. Further, not all feasibility studies are pilot studies, and may take on alternative

designs to the definitive study. However, all pilot studies constitute a type of feasibility study and researchers are encouraged to carefully reflect on both terms and use the recommended nomenclature with the addition of "a pilot study", "a pilot randomized controlled trial", "a feasibility study", etc. Readers are referred to an in-depth discussion about the accepted definitions and appropriateness of the different nomenclature [3].

3. Explicitly state the feasibility objectives and outcomes

Feasibility objectives along with the associated outcome measures must be preplanned (with rationale for and uncertainties of the definitive study in mind [19,29]), clearly stated, and defined in the methods section of the report. Definitions of feasibility outcomes should also include how and when they will be assessed [19]. This will help to prevent shortfalls such as the ones observed in several 'pilot' studies which concluded that there is a need to study feasibility (e.g., adherence, 'usefulness' to a group with more advanced disease) in the future [30,31]. Authors of one study mentioned that they carried out a previous evaluation of the study protocol to confirm acceptability and safety (although this was not published) [32]. Similar practices may be helpful to researchers for planning in the form of an initial acceptability or needs assessment survey, for example. A list of common feasibility outcomes that were searched or identified in this review are provided in Table 5, derived from previous research by the authors and other published work[1,2]. This list is not exhaustive and provided only as a guide.

4. Provide criteria for assessing the success of feasibility

Criteria for determining the success of feasibility, or progression criteria, should

be provided alongside explicit descriptions of feasibility outcomes. This also helps to avoid scenarios where some authors anecdotally comment on feasibility in discussions without formalizing feasibility or a priori feasibility outcomes. Although there have been previous calls for guidelines for anecdotal reports of more serious, health-related outcomes such as suspected adverse drug reactions [33], anecdotal reports of feasibility outcomes appear to be largely a result of the exploratory nature of some research and poor planning in other cases. Similarly, early stopping for 'efficacy' is an equally relevant issue in the context of pilot and feasibility studies as it is in full trials, since large treatment effects are more likely to be observed in small studies with small numbers of events [34,35] and may compromise actual success of the definitive study. Good examples of clearly stated feasibility criteria by authors of the included studies are provided in Table 4.

5. Provide justification for the sample size

For accurate inferences from pilot and feasibility study data, sampled participants need to be representative and based on the same eligibility criteria as the target population [1]. As such, the sample size for the pilot study needs to be justified as it would for the main study, in line with the feasibility objectives. One straightforward approach involves the calculation of a 95% CI for a proportion of participants based on a priori feasibility outcomes and criteria (e.g., 85% recruitment rate) [1]. Additional approaches, including those based on a desired level of precision around an estimated rate (e.g., recruitment), or a proportion of the expected sample size of the main study (e.g., 15% of what would be required), are outlined in the CONSORT pilot and feasibility trials extension [19].

Researchers should account for any uncertainty in the estimates where appropriate [19] and seek additional statistical expertise as needed to ensure best practices.

6. Interpret feasibility findings and report next steps for the definitive study

The discussion should focus on feasibility, implications for progression and proposed modifications (if any) to the design of the definitive study, with limitations addressing possible sources of bias and any uncertainty regarding feasibility [19]. Where possible, the point estimate of effect along with precision should be reported for feasibility outcomes [1]. Among the reviewed studies, many authors commented on a 'future larger study' but often in a more removed sense and from the perspective of sample size rather than feasibility for a future definitive study. In line with this, Kaur et al. found that among a sample of 191 pilot studies published in *Clinical Rehabilitation*, 85% were not followed-up with a definitive trial (and that this was not associated with the strength of the effect in the pilot study) [36]. The authors also state that due to the very large effect sizes (with small studies being more susceptible to this), studies may not progress as a result of beliefs that it is no longer ethical to offer the control [36]. Examples of adequate conclusions by authors of the included studies are outlined in Supplemental Table 1 (Supplemental Digital Content 3). Orsmond and Cohn also provide an overview of feasibility objectives and follow-up questions as a guide to understanding various barriers for success in studies focused on social and behavioural interventions [37].

7. Refer to and cite appropriate reporting guidelines

Complete reporting of research has a direct impact on its interpretation. The

extension to the 2010 CONSORT for randomized pilot and feasibility trials was published in 2016 [19], and is freely accessible as a guide to researchers for reporting randomized pilot and feasibility studies. This should supplement the use of a new reporting tool, the Randomized Controlled Trial Rehabilitation Checklist (RCTRACK), which has been developed to address an identified priority to improve the reporting of randomized trials in rehabilitation research [28,38]. Additionally, EQUATOR is the leading resource for reporting guidelines [39] which includes a library categorized by study design and clinical area. Additional reporting guidelines that may be relevant to rehabilitation pilot and feasibility studies include: STROBE (for case-control, cohort and cross-sectional designs) [40], TIDieR (for descriptions of interventions for replication) [41] with suggested considerations for rehabilitation research [42], the Transparent Reporting of Evaluation with Non-randomized Designs (TREND) statement (for behavioural and public health interventions, relevant for community-level interventions) [43], and the Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM) framework (for implementation research) [44]. Recommendations outlining amendments to the CONSORT pilot and feasibility trials extension for non-randomized designs (i.e., adapting items as necessary) have also been published [26].

Based on a sample of studies published in high impact rehabilitation journals, researchers found that only 14 of 139 (10.1%) randomized trials and observational studies declared the use of reporting guidelines (with just over half using them appropriately) [45]. Journal endorsement of reporting guidelines influences authors' reporting practices (e.g., fuller reporting) [46], and collaborative acknowledgment of this by editors [47] is encouraged.

However, Innocenti et al. concluded that such efforts have not impacted author declarations of reporting guideline use in rehabilitation research thus far [45]. This suggests that training of peer reviewers [45] and inclusion of reporting guideline methodological experts [48] is needed to support editorial efforts. The journal *Pilot and Feasibility Studies,* which focuses on promoting and publishing feasibility research, may be another alternative for researchers.

8. Report ethics, registration and funding information

Although missing ethics and funding information was not as pervasive an issue in the reviewed sample of studies, 10 (10.0%) studies did not provide a statement of ethics approval or exemption, and funding information was unclear or not reported in five (4.0%) studies. Statements mentioning that participants were enrolled or that informed consent was obtained "according to guidelines of the Institutional Review Board" are vague and should be avoided. Instead, researchers should explicitly state that the study was approved with the date of approval and approval number. Furthermore, only a quarter (23.0%) of included studies were registered. There are now well-established registries for trials including ClinicalTrials.gov (as of 2000), EudraCT (as of 2011), the International Standard Randomised Controlled Trial Number (ISRCTN, as of 2000) registry, as well as various national registries. Authors of pilot and feasibility studies should follow the same standard practices as relevant to full trials, including pre-registration, to prevent post hoc decision-making and selective reporting.

Limitations

The findings in this review should be interpreted with caution. Previously described inferences about feasibility implications and conclusions in the included studies were based on investigator judgments. No statistical tests were performed to investigate associations with design and reporting practices, as this was beyond the scope of the paper. For generalizability, a random sample of studies was selected but reporting practices vary over time with the development of guidelines and registries. Potential temporal trends in publishing were unaccounted for as the sample of studies was not stratified by year. Future assessments of reporting completeness should be evaluated against the publication of reporting guidelines.

CONCLUSIONS

This review has highlighted that pilot and feasibility studies are poorly defined, designed and reported in rehabilitation research. This contributes to wastefulness in research and resources, compromises education efforts, and minimizes patient contributions to research. This paper summarizes key considerations for researchers who are interested in designing pilot and feasibility studies to encourage conscientious, preplanned methodology. This will also help to prevent inadequate research practices such as post hoc labeling of small studies as pilots, anecdotal conclusions about feasibility, and underpowered analyses to determine treatment effects in studies intended to evaluate the feasibility of future, definitive research.

ETHICS

Ethics committee approval and consent to participate was not required for this research since only previously published data was used.

AUTHORS' CONTRIBUTIONS

LT, DOL and LM conceived the idea and contributed to the design of the study. DOL, KM and SE screened the studies. DOL, KM, SE, CL, KHK, and KK extracted the data. DOL analyzed the data and wrote the first draft of the manuscript. DOL, LM and LT interpreted the data, and LM and LT provided critical feedback on the manuscript. All authors read and approved the final version.

ACKNOWLEDGEMENTS

We are indebted to Dr. Dinesh Kumbhare for the invitation to carry out this research and for his guidance in the development of the study protocol. We would like to thank Denise Smith at the Health Sciences Library at McMaster University for her support in developing the review search strategy.

REFERENCES

1. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol.* 2010;10:1.
2. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. *Am J Prev Med.* 2009;36(5):452-457.
3. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. *PLoS One.* 2016;11(3):e0150205.
4. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. *J Eval Clin Pract.* 2004;10(2):307-312.
5. Ioannidis JP, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet.* 2014;383(9912):166-175.
6. Pilot studies: common uses and misuses. National Center for Complementary and Integrative Health. https://www.nccih.nih.gov/grants/pilot-studies-common-uses-and-misuses. Accessed March 4, 2021.
7. Fuhrer MJ. Overview of clinical trials in medical rehabilitation: impetuses, challenges, and needed future directions. *Am J Phys Med Rehabil.* 2003;82(10 Suppl):S8-15.
8. El-Kotob R, Giangregorio LM. Pilot and feasibility studies in exercise, physical activity, or rehabilitation research. *Pilot Feasibility Stud.* 2018;4:137.
9. Kho ME, Thabane L. Pilot and Feasibility Studies in Rehabilitation: Moving into the Next Decade. *Physiotherapy Canada.* 2020;72(3):225-227.

10. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Med Res Methodol.* 2010;10:67.

11. Charlesworth M, Klein AA, White SM. A bibliometric analysis of the conversion and reporting of pilot studies published in six anaesthesia journals. *Anaesthesia.* 2020;75(2):247-253.

12. Shanyinde M, Pickering RM, Weatherall M. Questions asked and answered in pilot and feasibility randomized controlled trials. *BMC Med Res Methodol.* 2011;11:117.

13. Jones TA, Olds TS, Currow DC, Williams MT. Feasibility and Pilot Studies in Palliative Care Research: A Systematic Review. *J Pain Symptom Manage.* 2017;54(1):139-151 e134.

14. Mbuagbaw L, Kosa SD, Lawson DO, et al. The reporting of progression criteria in protocols of pilot trials designed to assess the feasibility of main trials is insufficient: a meta-epidemiological study. *Pilot Feasibility Stud.* 2019;5:120.

15. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev.* 2016;5(1):210.

16. Covidence systematic review software. Veritas Health Innovation. www.covidence.org. Published 2020. Accessed January 8, 2021.

17. Mbuagbaw L, Lawson DO, Puljak L, Allison DB, Thabane L. A tutorial on methodological studies: the what, when, how and why. *BMC Med Res Methodol.* 2020;20(1):226.

18. *Microsoft Excel for Mac* [computer program]. Version 16.46. Redmond, WA: Microsoft; 2021.

19. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ.* 2016;355:i5239.

20. *IBM SPSS Statistics for Macintosh* [computer program]. Version 24.0. Armonk, NY: IBM Corp.; 2016.

21. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ.* 1996;313(7057):570-571.

22. Gandolfi M, Geroin C, Vale N, et al. Does myofascial and trigger point treatment reduce pain and analgesic intake in patients undergoing onabotulinumtoxinA injection due to chronic intractable migraine? *European journal of physical and rehabilitation medicine.* 2018;54(1):1-12.

23. Avery KN, Williamson PR, Gamble C, et al. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. *BMJ Open.* 2017;7(2):e013537.

24. Resources for pilot and feasibility studies. Queen Mary University of London. https://pilotandfeasibilitystudies.qmul.ac.uk/resources/. Accessed March 8, 2021.

25. Trial Forge: Feasibility & pilot work. University of Aberdeen. https://www.trialforge.org/pathway/feasibility-and-pilot-work/. Accessed March 8, 2021.

26. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. *Pilot Feasibility Stud.* 2019;5:114.

27. Goodacre S. Uncontrolled before-after studies: discouraged by Cochrane and the EMJ. *Emerg Med J.* 2015;32(7):507-508.

28. Levack WM, Meyer T, Negrini S, Malmivaara A. Cochrane Rehabilitation Methodology Committee: an international survey of priorities for future work. *Eur J Phys Rehabil Med.* 2017;53(5):814-817.

29. Campbell MJ, Mansournia MA, Lancaster G. Methods matter: pilot and feasibility studies in sports medicine. *Br J Sports Med.* 2020;54(22):1309-1310.

30. Akinwuntan AE, Devos H, Baker K, et al. Improvement of driving skills in persons with relapsing-remitting multiple sclerosis: A pilot study. *Archives of Physical Medicine and Rehabilitation.* 2014;95(3):531-537.

31. Atan T, Karavelioglu Y. Effectiveness of high-intensity interval training versus moderate-intensity continuous training in patients with fibromyalgia: A pilot randomized controlled trial. *Archives of physical medicine and rehabilitation.* 2020;22.

32. Iwatsu K, Yamada S, Iida Y, et al. Feasibility of neuromuscular electrical stimulation immediately after cardiovascular surgery. *Arch Phys Med Rehabil.* 2015;96(1):63-68.

33. Aronson JK. Anecdotes as evidence. *BMJ.* 2003;326(7403):1346.

34. Bassler D, Montori VM, Briel M, Glasziou P, Guyatt G. Early stopping of randomized clinical trials for overt efficacy is problematic. *J Clin Epidemiol.* 2008;61(3):241-246.

35. Montori VM, Devereaux PJ, Adhikari NK, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA.* 2005;294(17):2203-2209.

36. Kaur N, Figueiredo S, Bouchard V, Moriello C, Mayo N. Where have all the pilot studies gone? A follow-up on 30 years of pilot studies in Clinical Rehabilitation. *Clin Rehabil.* 2017;31(9):1238-1248.

37. Orsmond GI, Cohn ES. The Distinctive Features of a Feasibility Study: Objectives and Guiding Questions. *OTJR (Thorofare N J).* 2015;35(3):169-177.

38. Negrini S, Armijo-Olivo S, Patrini M, et al. The Randomized Controlled Trials Rehabilitation Checklist: Methodology of Development of a Reporting Guideline Specific to Rehabilitation. *Am J Phys Med Rehabil.* 2020;99(3):210-215.

39. Enhancing the QUAlity and Transparency Of health Research (EQUATOR). University of Oxford. https://www.equator-network.org. Accessed January 29, 2021.

40. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol.* 2008;61(4):344-349.

41. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ.* 2014;348:g1687.

42. Whyte J, Dijkers MP, Fasoli SE, et al. Recommendations for Reporting on Rehabilitation Interventions. *Am J Phys Med Rehabil.* 2021;100(1):5-16.

43. Des Jarlais DC, Lyles C, Crepaz N, Group T. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health.* 2004;94(3):361-366.

44. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health.* 1999;89(9):1322-1327.

45. Innocenti T, Salvioli S, Giagio S, Feller D, Cartabellotta N, Chiarotto A. Declaration of use and appropriate use of reporting guidelines in high-impact rehabilitation journals is limited: a meta-research study. *J Clin Epidemiol.* 2020;131:43-50.

46. Turner L, Shamseer L, Altman DG, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev.* 2012;11:MR000030.

47. Chan L, Heinemann AW, Roberts J. Elevating the quality of disability and rehabilitation research: mandatory use of the reporting guidelines. *Eur J Phys Rehabil Med.* 2014;50(2):125-127.

48. Blanco D, Schroter S, Aldcroft A, et al. Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial. *BMJ Open.* 2020;10(5):e036799.

49. Morone G, Pisotta I, Pichiorri F, et al. Proof of principle of a brain-computer interface approach to support poststroke arm rehabilitation in hospitalized

patients: Design, acceptability, and usability. *Archives of Physical Medicine and Rehabilitation.* 2015(3):S71-S78.

50. Bourke L, Thompson G, Gibson DJ, et al. Pragmatic lifestyle intervention in patients recovering from colon cancer: A randomized controlled pilot study. *Archives of Physical Medicine and Rehabilitation.* 2011;92(5):749-755.

51. Moseley Jr JB, Wray NP, Kuykendall D, Willis K, Landon G. Arthroscopic treatment of osteoarthritis of the knee: A prospective, randomized, placebo-controlled trial: Results of a pilot study. *American Journal of Sports Medicine.* 1996;24(1):28-34.

FIGURE LEGENDS

Figure 1. Flow diagram for the review.

Figure 2. Pyramid graph of proportion (percentages) of studies reporting a primary objective related to feasibility as compared to studies that did not report a primary feasibility objective.

TABLES

Table 1. Eligibility criteria for the review.

Table 2. General characteristics of included studies.

Table 3. Estimates of good practices related to feasibility in included studies.

Table 4. Overview with illustrative examples of feasibility outcomes, associated progression criteria and results as reported by authors of the included studies.

Table 5. Examples of feasibility outcomes which may be relevant for rehabilitation researchers.

SUPPLEMENTAL DIGITAL CONTENT

Supplemental Digital Content 1: search strategy

Supplemental Digital Content 2: list of included studies

Supplemental Digital Content 3: overview with illustrative examples of feasibility conclusions as reported by authors of the included studies (Supplemental Table 1)

Figure 1

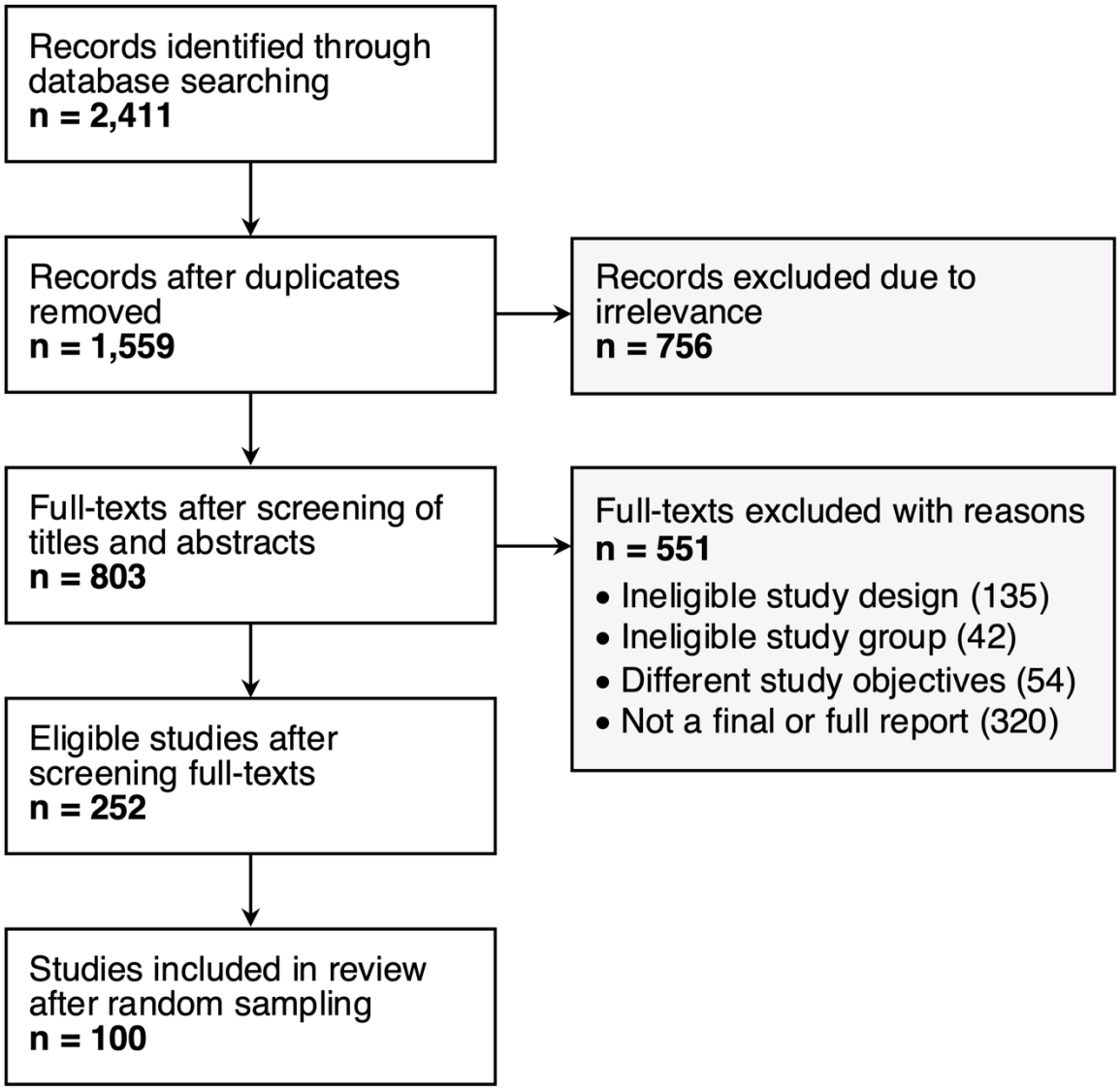

Records identified through database searching
n = 2,411
Records after duplicates removed
n = 1,559
Records excluded due to irrelevance
n = 756
Full-texts after screening of titles and abstracts
n = 803
Full-texts excluded with reasons
n = 551
• Ineligible study design (135)
• Ineligible study group (42)
• Different study objectives (54)
• Not a final or full report (320)
Eligible studies after screening full-texts
n = 252
Studies included in review after random sampling
n = 100

Figure 2

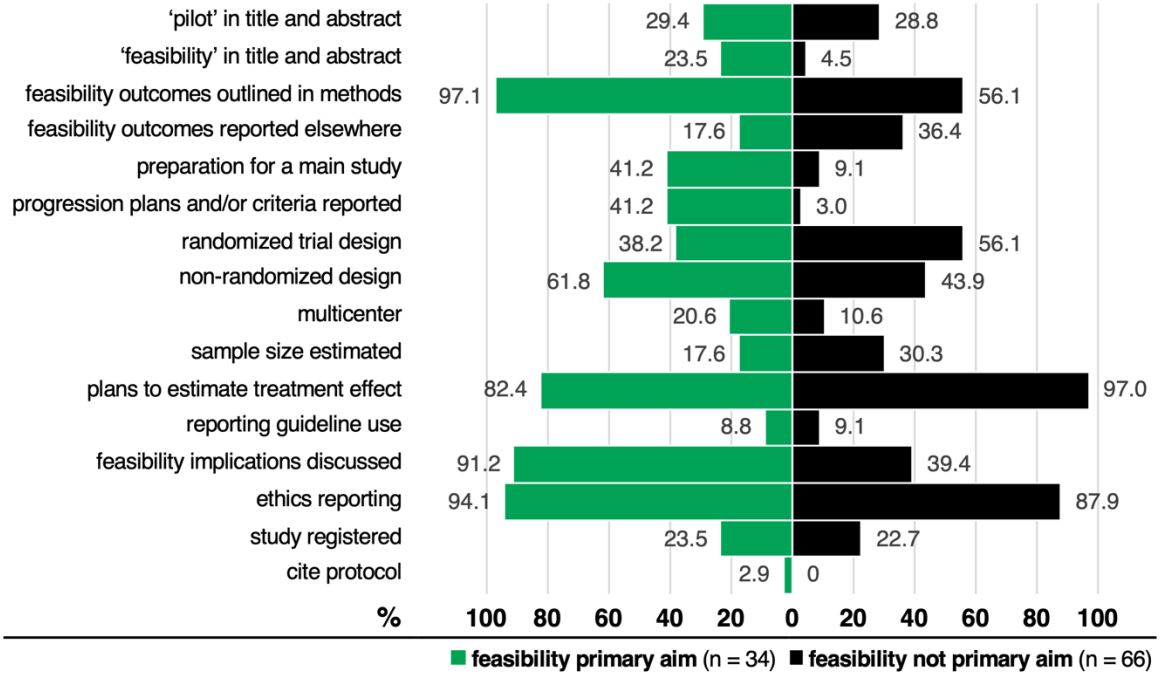

'pilot' in title and abstract
'feasibility' in title and abstract
feasibility outcomes outlined in methods
feasibility outcomes reported elsewhere
preparation for a main study
progression plans and/or criteria reported
randomized trial design
non-randomized design
multicenter
sample size estimated
plans to estimate treatment effect
reporting guideline use
feasibility implications discussed
ethics reporting
study registered
cite protocol
29.4
28.8
23.5
4.5
97.1
56.1
17.6
36.4
41.2
9.1
41.2
3.0
38.2
56.1
61.8
43.9
20.6
10.6
17.6
30.3
82.4
97.0
8.8
9.1
91.2
39.4
94.1
87.9
23.5
22.7
2.9
0
%
100 80 60 40 20 0 20 40 60 80 100
feasibility primary aim (n = 34)
feasibility not primary aim (n = 66)

Table 1. Eligibility criteria for the review.

Included	**Excluded**
• Any article where the authors used the terms "pilot" or" feasibility" in the title or abstract in relation to the study design • Studies where the aim of the research was focused on health outcomes in patients and their families/caregivers • Full-text, final reports of studies (i.e., to allow for an in-depth evaluation of study design and reporting practices)	• Studies where the primary group of interest was a non-patient group (e.g., healthy population, health professionals) • Other types of study reports (e.g., conference abstracts, brief reports, letters and commentaries, protocols, reviews) • Other study designs: - retrospective studies including retrospective analyses of studies - analytical or methods studies (e.g., assessing feasibility of a predictive model, or new method of isolating cells for therapy) - exploratory studies (e.g., assessing factors associated with a condition) - studies assessing feasibility in the context of a psychometric, reliability, or validation study (e.g., development of a tool, core outcome sets) - case reports and case series

Table 2. General characteristics of included studies.

Characteristic	***n (%)***	***95% CI***
Total	100	—
Journal		
American Journal of Physical Medicine and	14 (14.0)	8.3-21.8
Rehabilitation	4 (4.0)	1.4-9.2
American Journal of Sports Medicine	52 (52.0)	42.3-61.6
Archives of Physical Medicine and Rehabilitation	8 (8.0)	3.9-14.5
British Journal of Sports Medicine	10 (10.0)	5.3-17.0
European Journal of Physical and Rehabilitation	12 (12.0)	6.7-19.4
Medicine		
PM&R Journal		
Year of publication		
1975 – 84	1 (1.0)	0.1-4.6
1985 – 94	1 (1.0)	0.1-4.6
1995 – 04	11 (11.0)	6.0-18.2
2005 – 14	47 (47.0)	37.4-56.8
2015 – 20	40 (40.0)	30.8-49.8
Region [a]		
Asia	9 (8.9)	4.5-15.8
Europe	27 (26.7)	19.0-36.3
North America	54 (53.5)	44.2-63.6
Oceania	6 (5.9)	2.5-12.0
Other	5 (5.0)	1.9-10.6
Overarching methodological approach		
Quantitative	97 (97.0)	92.2-99.1
Qualitative and quantitative	3 (3.0)	0.9-7.8
Study design [b]		
Randomized trial	50 (49.5)	40.3-59.7
Cohort (prospective)	1 (1.0)	0.1-4.6
Controlled before-after	7 (6.9)	3.2-13.3
Crossover (non-randomized)	1 (1.0)	0.1-4.6
Uncontrolled before-after	41 (40.6)	31.7-50.8
Focus group	1 (1.0)	0.1-4.6
Centers involved		
Single	85 (85.0)	77.0-91.0
Multiple	14 (14.0)	8.3-21.8
Not reported	1 (1.0)	0.1-4.6
Cite a published protocol (yes)	1 (1.0)	0.1-4.6
Study registered (yes)	23 (23.0)	15.6-31.9
Condition or disease investigated*		
Aging-related/geriatric	6 (5.3)	2.5-12.0

Arthritis	9 (8.0)	4.5-15.9
Cancer	6 (5.3)	2.5-12.0
Cardiovascular	5 (4.4)	1.9-10.6
Chronic (non-cancer) pain	8 (7.1)	3.8-14.6
Fractures or soft tissue injuries	9 (8.0)	4.5-15.9
Movement or neuromuscular disorders (e.g., Parkinson's)	10 (8.8)	5.2-17.1
Spinal cord injury	11 (9.7)	6.0-18.3
Stroke	23 (20.4)	15.5-32.2
Other	26 (23.0)	18.1-35.5
Age of eligible participants [c]		
Less than 18 years	6 (5.9)	2.5-12.0
18 years and older	52 (51.0)	42.2-61.8
65 years and older	7 (6.9)	3.2-13.3
Not reported	37 (36.3)	28.0-46.8
Number of patient participants analyzed, mean (SD, range)	28.9 (29.4, 3-194)	23.1-34.8
Non-patient participants included*		
Caregivers	1 (1.0)	0.1-4.6
Clinicians	3 (3.0)	0.9-7.8
Study personnel (e.g., research coordinators)	1 (1.0)	0.1-4.6
None	96 (95.0)	90.4-99.1
Type of experimental intervention		
Injection	6 (6.0)	2.5-11.9
Medication or pharmacologic-based	6 (6.0)	2.5-11.9
Movement or exercise-based	48 (48.0)	38.4-57.7
Physical modality	15 (15.0)	9.0-23.0
Occupational	4 (4.0)	1.4-9.2
Other or multi-modal	21 (21.0)	13.9-29.7
Type of comparator intervention		
Standard of care	28 (28.0)	19.9-37.3
Placebo/sham	11 (11.0)	6.0-18.2
Waitlist/wait-and-see	5 (5.0)	1.9-10.6
Movement or exercise-based	10 (10.0)	5.3-17.0
Other or multi-modal	3 (3.0)	0.9-7.8
Not applicable (single arm study)	43 (43.0)	33.6-52.8
Performed intention-to-treat analysis		
Yes	9 (9.0)	4.5-15.8
No	34 (34.0)	25.3-43.6
Partially	4 (4.0)	1.4-9.2
Unclear	3 (3.0)	0.9-7.8
Not applicable (not randomized trial)	50 (50.0)	40.3-59.7
Research ethics approval/waiver statement (yes)	90 (90.0)	83.0-94.7
Funding type*		

Government	34 (27.2)	25.0-44.3
Institutional (e.g., hospitals, universities)	23 (18.4)	15.5-32.3
Industry	9 (7.2)	4.5-15.9
Non-profit (e.g., charities, foundations)	25 (20.0)	17.1-34.6
Other (e.g., facilities, professional societies)	5 (4.0)	1.9-10.7
None	24 (19.2)	16.3-33.5
Not reported	5 (4.0)	1.9-10.7

CI: confidence interval for percentage, n: number of articles, SD: standard deviation

* studies can be counted towards more than one category

[a] one study conducted in Germany (Europe) and Japan (Asia)

[b] one study included a prospective cohort and focus group

[c] two studies included patients less than 18 years and older

Table 3. Estimates of good practices related to feasibility in included studies.

Characteristic	***n* (%)**	***95% CI***
Total	100	—
Naming conventions		
'Pilot' in title	30 (30.0)	21.7-39.5
'Pilot' in abstract	17 (17.0)	10.6-25.2
'Pilot' in both	29 (29.0)	20.8-38.4
'Feasibility' in title	3 (3.0)	0.9-7.8
'Feasibility' in abstract	23 (23.0)	15.6-31.9
'Feasibility' in both	11 (11.0)	6.0-18.2
Methods		
Primary objective related to study feasibility (yes)	34 (34.0)	25.3-43.6
Feasibility outcomes are provided (yes)	70 (70.0)	60.5-78.3
Feasibility outcomes investigated*		
Acceptance (e.g., adoption, comfort, satisfaction,	32 (15.6)	22.8-43.1
tolerance)	1 (0.5)	0.1-4.6
Adherence/compliance with protocol (clinician)	40 (19.5)	29.8-52.0
Adherence/compliance with protocol (patient)	2 (1.0)	0.4-6.3
Data completeness rate	9 (4.4)	4.5-16.1
Data for sample size estimate	7 (3.4)	3.2-13.5
Enrollment/recruitment rate	6 (2.9)	2.5-12.2
Feedback on intervention (clinician)	23 (11.2)	15.3-33.0
Feedback on intervention (patient)	3 (1.5)	0.8-7.9
Logistics of, or resources required to conduct definitive	2 (1.0)	0.4-6.3
study	12 (5.9)	6.6-19.9
Preliminary efficacy	36 (17.6)	26.3-47.6
Participation/retention rate (e.g., attrition, complete	2 (1.0)	0.4-6.3
follow-ups)	30 (14.6)	21.1-40.9
Safety		
Other		
Not applicable (none provided)		
Sample size estimated (yes)	25 (25.0)	17.3-34.1
Sample size estimation justification*		
Adequate: Feasibility outcomes	1 (0.9)	0.1-4.6
Proportion of a larger trial	0 (0)	—
Recommendations in literature/rule of thumb	0 (0)	—
Inadequate: Intervention effect	18 (15.8)	11.4-26.6
Other similar studies	19 (16.7)	12.2-27.7
Resource availability	1 (0.9)	0.1-4.6
None	0 (0)	—
Not applicable (no estimation)	75 (65.8)	64.7-84.4

Declaration that study done in preparation for a main study (yes)	21 (21.0)	13.9-29.7
Progression plans and criteria provided		
Progression plans mentioned only	14 (14.0)	8.3-21.8
Progression plans with criteria for ≥ one feasibility outcome	4 (4.0)	1.4-9.2
No	82 (82.0)	73.6-88.6
Intent for hypothesis testing or to provide an estimate of treatment effect (yes)	92 (92.0)	85.5-96.1
Reporting guideline referenced*		
CONSORT, generic version for RCTs	5 (5.0)	1.9-10.6
CONSORT, non-pharmacologic treatment RCTs extension	1 (1.0)	0.1-4.6
CONSORT, pilot and feasibility RCTs extension	2 (2.0)	0.4-6.3
STROBE	1 (1.0)	0.1-4.6
TIDieR	1 (1.0)	0.1-4.6
Not applicable (no reference)	91 (90.1)	84.0-95.7
Journal reporting guidance level		
High to moderate	92 (92.0)	85.5-96.1
Low to none	8 (8.0)	3.9-14.5
Results		
Additional feasibility outcomes are provided† (yes)	30 (30.0)	21.7-39.5
Additional feasibility outcomes as reported in Results/Discussion sections only*		
Acceptance (e.g., adoption, comfort, satisfaction, tolerance)	8 (7.0)	3.8-14.6
Adherence/compliance with protocol (clinician)	1 (0.9)	0.1-4.6
Adherence/compliance with protocol (patient)	12 (10.4)	6.7-19.5
Enrollment/recruitment rate	4 (3.5)	1.4-9.3
Feedback on intervention (patient)	6 (5.2)	2.5-12.0
Logistics of, or resources required to conduct definitive study	1 (0.9)	0.1-4.6
Participation/retention rate (e.g., attrition, complete follow-ups)	4 (3.5)	1.4-9.3
Safety	9 (7.8)	4.5-15.9
Not applicable (none provided)	70 (60.9)	59.5-79.8
Discussion		
Implications of feasibility findings on a main study discussed (yes)	57 (57.0)	47.2-66.4
Conclusions with regard to feasibility of a main study		
Stop – Not feasible	0 (0)	—
Caution – Feasible with modifications	43 (43.0)	33.6-52.8
Start – Feasible as is	14 (14.0)	8.3-21.8

Not discussed	43 (43.0)	33.6-52.8

CI: confidence interval for percentage, CONSORT: Consolidated Standards of Reporting Trials, n: number of articles, n: number of articles, PAFS: pilot and feasibility studies, RCT: randomized controlled trials, STROBE: Strengthening the Reporting of Observational Studies in Epidemiology, TIDieR: Template for Intervention Description and Replication

* studies can be counted towards more than one category

† includes n = 4 studies that commented on feasibility anecdotally

Table 4. Overview with illustrative examples of feasibility outcomes, associated progression criteria and results as reported by authors of the included studies.

Scenario	**Example(s) of feasibility outcomes investigated and criteria for progression**	**Feasibility results**
Progression plans only		
a) Brain-computer interface-assisted motor imagery training for stroke (cohort study with focus group) [49]	• Feedback on intervention (clinician)	• "High learnability emerged in the focus group therapists" and "The acceptability of the proposed system by therapists highly depends on a subjective technical confidence and a positive attitude toward the use of technologies." Based on focus group feedback, authors state that modifications are needed to address the time required to set up the system.
b) Pragmatic lifestyle intervention for colon cancer (randomized trial) [50]	• Data for sample size estimate	• "Therefore, using the FACT-C data from our intervention group (mean ± SD, 120 ± 15), the required sample size to show a clinically important effect is estimated to be 57 participants a group. Given a 6% dropout rate in the present study over the course of the intervention, this would require a cohort of 122 (61 patients in each group) to detect significance at an alpha level of .05 with 80% power."
Progression plans with criteria		
c) Placebo effect in arthroscopy for osteoarthritis (randomized trial) [51]	1. Enrollment/recruitment rate, criteria: ability to find and recruit eligible patients 2. Other – maintenance of blinding (patient, evaluator)	1. "Four hundred thirty-five patients presented with knee problems and, of these, 241 (55%) had radiographic evidence of osteoarthritis. Fifty-five (23%) of the 241 patients with osteoarthritis met all of the eligibility criteria. Of these 55

	3. Participation/retention rate, criteria: ability to retain patients for at least 6 months	patients, 26 (47%) agreed to participate in the study. The recruitment rate during the pilot phase was therefore 1.63 patients/week (26 patients in 16 weeks). The 26 patients who agreed to enroll were in addition to the 10 patients who participated in the pilot study," (criteria met) 2. "Thus, at each postoperative visit the patients and examining physician were questioned as to which procedure they believed the patient had received. The patients and the physician were never able to guess the procedure correctly more often than would be explained by chance alone." (criteria met) 3. Of the ten enrolled and randomized patients, "[one] subject was dropped from the study at the 3-month postoperative followup because of an acute psychiatric episode... We have since revised our eligibility criteria to reject subjects who have a history of psychiatric illness. All nine eligible subjects participated in the 6-month evaluation." (criteria partially met)
d) Neuromuscular electrical stimulation after cardiovascular surgery (uncontrolled before-after study) [32]	• Authors explicitly predefined all criteria for feasibility with justifications based on clinically acceptable ranges (e.g., clinical expert opinion, literature): 1. Adherence/compliance with protocol (patient), criteria: >80% of	1. "Sixty-one (89.7%) of 68 patients completed the NMES intervention." (criteria met) 2. "No patients had changes in blood pressure and heart rate during the study period that exceeded the criteria values." (criteria met) 3. "No adverse effects of NMES on temporary epicardial pacing" (criteria met)

	patients able to complete more than 4 of 5 sessions 2. Safety, criteria: systolic blood pressure >80mmHg and change in systolic blood pressure <20mmHg 3. Safety, criteria: no malfunction of temporary epicardial pacing during session 4. Safety, criteria: incidence of new onset of postoperative AF during the study period <30% in coronary artery bypass surgery, <40% in valvular surgery, and <50% in combined and aortic surgery	4. "The incidence of postoperative AF during the study period was 26.9% (7/26) in coronary artery bypass surgery, 18.2 % (4/22) in valvular surgery, and 20.0% (4/20) in aortic or combined surgery." (criteria met)

AF: atrial fibrillation, FACT-C: Functional Assessment of Cancer Therapy-Colorectal cancer, NMES: neuromuscular electrical stimulation, SD: standard deviation

Table 5. Examples of feasibility outcomes which may be relevant for rehabilitation researchers.

Outcome	Interpretation and approaches to measurement
Acceptance	• Acceptability of the intervention by participants and/or administrators of the intervention • Surrogates for acceptance include: - adoption of practices (e.g., reports of continued use after study) - comfort and fit of devices - degree of exertion or workload required - recommendations to others (e.g., family members with similar condition) - tolerance - satisfaction
Adherence	• (Non)compliance to the intervention protocol and procedures • May include measures of the degree of consistency (e.g., intraclass correlation coefficient) between and within administrators of the intervention
Contamination	• Study design features or processes that lead to control arm participants' exposure to the intervention (e.g., elicited through verbal queries or questionnaires)
Data collection processes and completeness	• Ability to capture all relevant study data • Reliability of data collection tools or personnel (e.g., identification of gaps that lead to missing data, training needs or calibration requirements)
Dropouts or retention	• Proportion of participants that are lost to follow-up, or remain in the study (e.g., number of complete follow-ups)
Enrollment or recruitment	• Proportion of participants that agree (or refuse) to participate
Feedback on intervention	• General comments or suggestions for improvement to the intervention (e.g., collected with scales, questionnaires or open-ended feedback forms, verbally during follow-up calls or visits) • May include participants receiving intervention, caregivers and family members, administrators of the intervention (e.g., ease of implementation, willingness to incorporate into routine practice), research personnel, etc.
Logistics, management and required resources	• Organizational aspects and capacity for implementation of the intervention: - integration with existing structures, spaces and workflow

	- implementation without significant modifications or disruptions to existing structures, spaces and workflow - availability of equipment, software, facility or storage space, personnel - personnel training requirements, or degree of support required - associated costs
Participation	• Similar to adherence, takes into account participant's choice especially when intervention components are optional (e.g., number of sessions attended) • May be more relevant for pragmatic or community-based interventions
Potential efficacy, effectiveness, or data for sample size calculations	• Preliminary estimates of treatment effect and variance, based on relevant clinical outcomes as defined by investigators • May be used to generate estimates for the required sample size in the definitive study
Safety	• Measures such as rates of adverse events or reactions • May be used to determine appropriate dosages of intervention
Study design considerations	• Determination of the appropriateness of various components of an intervention: - appropriateness of eligibility criteria, clinical outcome measures, tests, etc. - need for incorporating or removing certain features (e.g., study arm) - need for developing or adapting clinical outcome measures or tests for improved accessibility, legibility/health literacy levels, etc. - validity of clinical outcome measures (e.g., face validity, differences between subjective and objective measures for answering the study question) - successful maintenance of design features (e.g., blinding of participants and evaluators throughout the study duration)
Timeliness	• Time required to complete the intervention protocol and procedures by participants (e.g., to fill out questionnaire) or by administrators of the intervention (e.g., to perform injection)

As cited in

"Pilot and feasibility studies in rehabilitation research: a review and educational primer for the physiatrist researcher"

SUPPLEMENTAL DIGITAL CONTENT 1

Search strategy (searched on July 31, 2020)

Medline via Ovid:

Concept: top 6 journals in rehabilitation research		
1	"american journal of physical medicine & rehabilitation".jn. OR "american journal of sports medicine".jn. OR "archives of physical medicine & rehabilitation".jn. OR "british journal of sports medicine".jn. OR "european journal of physical & rehabilitation medicine".jn. OR "pm & r".jn.	38,477
Concept: pilot and feasibility studies		
2	Pilot Projects/	122,628
3	Feasibility Studies/	68,158
4	(pilot or feasibility).mp.	379,450
5	or/2-4	379,450
6	1 and 5	999

Embase via Ovid:

Concept: top 6 journals in rehabilitation research		
1	"american journal of physical medicine and rehabilitation".jn. OR "american journal of physical medicine rehabilitation" OR "american journal of physical medicine rehabilitation association of academic physiatrists" OR "american journal of sports medicine".jn. OR "archives of physical medicine and rehabilitation".jn. OR "british journal of sports medicine".jn. OR "european journal of physical and rehabilitation medicine".jn. OR "pm and r".jn. OR "pm r the journal of injury function and rehabilitation".jn.	40,423
Concept: pilot and feasibility studies		
2	pilot study/	155,266
3	feasibility study/	132,472
4	(pilot or feasibility).mp.	524,004
5	or/2-4	524,004
6	1 and 5	1,412

SUPPLEMENTAL DIGITAL CONTENT 2

List of included studies

1. Akinwuntan AE, Devos H, Baker K, et al. Improvement of driving skills in persons with relapsing-remitting multiple sclerosis: A pilot study. *Archives of Physical Medicine and Rehabilitation.* 2014;95(3):531-537.
2. Albert HB, Manniche C, Sorensen JS, Deleuran BW. Antibiotic treatment in patients with low-back pain associated with Modic changes Type 1 (bone oedema): A pilot study. *British Journal of Sports Medicine.* 2008;42(12):969-973.
3. Andrysek J, Klejman S, Steinnagel B, et al. Preliminary evaluation of a commercially available videogame system as an adjunct therapeutic intervention for improving balance among children and adolescents with lower limb amputations. *Archives of Physical Medicine and Rehabilitation.* 2012;93(2):358-366.
4. Atan T, Karavelioglu Y. Effectiveness of high-intensity interval training versus moderate-intensity continuous training in patients with fibromyalgia: A pilot randomized controlled trial. *Archives of physical medicine and rehabilitation.* 2020;22.
5. Athukorala RP, Jones RD, Sella O, Huckabee ML. Skill training for swallowing rehabilitation in patients with Parkinson's disease. *Arch Phys Med Rehabil.* 2014;95(7):1374-1382.
6. Backhaus SL, Ibarra SL, Klyce D, Trexler LE, Malec JF. Brain Injury Coping Skills Group: A Preventative Intervention for Patients With Brain Injury and Their Caregivers. *Archives of Physical Medicine and Rehabilitation.* 2010;91(6):840-848.
7. Ballaz L, Fusco N, Cretual A, Langella B, Brissot R. Peripheral Vascular Changes After Home-Based Passive Leg Cycle Exercise Training in People With Paraplegia: A Pilot Study. *Archives of Physical Medicine and Rehabilitation.* 2008;89(11):2162-2166.
8. Bauer P, Krewer C, Golaszewski S, Koenig E, Muller F. Functional electrical stimulation-assisted active cycling--therapeutic effects in patients with hemiparesis from 7 days to 6 months after stroke: a randomized controlled pilot study. *Arch Phys Med Rehabil.* 2015;96(2):188-196.
9. Bean JF, Brown L, DeAngelis TR, et al. The Rehabilitation Enhancing Aging Through Connected Health Prehabilitation Trial. *Arch Phys Med Rehabil.* 2019;100(11):1999-2005.
10. Boon AJ, Smith J, Dahm DL, et al. Efficacy of intra-articular botulinum toxin type A in painful knee osteoarthritis: a pilot study. *PM R.* 2010;2(4):268-276.
11. Bourke L, Thompson G, Gibson DJ, et al. Pragmatic lifestyle intervention in patients recovering from colon cancer: A randomized controlled pilot study. *Archives of Physical Medicine and Rehabilitation.* 2011;92(5):749-755.
12. Bruns ERJ, Argillander TE, Schuijt HJ, et al. Fit4SurgeryTV At-home Prehabilitation for Frail Older Patients Planned for Colorectal Cancer Surgery: A

Pilot Study. *American journal of physical medicine & rehabilitation.* 2019;98(5):399-406.
13. Caliandro P, Celletti C, Padua L, et al. Focal muscle vibration in the treatment of upper limb spasticity: A pilot randomized controlled trial in patients with chronic stroke. *Archives of Physical Medicine and Rehabilitation.* 2012;93(9):1656-1661.
14. Cannell LJ, Taunton JE, Clement DB, Smith C, Khan KM. A randomised clinical trial of the efficacy of drop squats or leg extension/leg curl exercises to treat clinically diagnosed jumper's knee in athletes: Pilot study. *British Journal of Sports Medicine.* 2001;35(1):60-64.
15. Carroll LM, Volpe D, Morris ME, Saunders J, Clifford AM. Aquatic Exercise Therapy for People With Parkinson Disease: A Randomized Controlled Trial. *Arch Phys Med Rehabil.* 2017;98(4):631-638.
16. Cebicci MA, Sutbeyaz ST, Goksu SS, Hocaoglu S, Oguz A, Atilabey A. Extracorporeal Shock Wave Therapy for Breast Cancer-Related Lymphedema: A Pilot Study. *Arch Phys Med Rehabil.* 2016;97(9):1520-1525.
17. Chen K, Wu YN, Ren Y, et al. Home-Based Versus Laboratory-Based Robotic Ankle Training for Children With Cerebral Palsy: A Pilot Randomized Comparative Trial. *Arch Phys Med Rehabil.* 2016;97(8):1237-1243.
18. Cheville AL, Girardi J, Clark MM, et al. Therapeutic exercise during outpatient radiation therapy for advanced cancer: Feasibility and impact on physical well-being. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2010;89(8):611-619.
19. Chieffo R, De Prezzo S, Houdayer E, et al. Deep repetitive transcranial magnetic stimulation with H-coil on lower limb motor function in chronic stroke: a pilot study. *Arch Phys Med Rehabil.* 2014;95(6):1141-1147.
20. Christiansen CL, Miller MJ, Murray AM, et al. Behavior-Change Intervention Targeting Physical Function, Walking, and Disability After Dysvascular Amputation: A Randomized Controlled Pilot Trial. *Archives of Physical Medicine and Rehabilitation.* 2018;99(11):2160-2167.
21. Chu VWT, Hornby TG, Schmit BD. Effect of antispastic drugs on motor reflexes and voluntary muscle contraction in incomplete spinal cord injury. *Archives of Physical Medicine and Rehabilitation.* 2014;95(4):622-632.
22. DePalma MJ, Ketchum JM, Queler ED, Trussell BS. Prospective pilot study of painful lumbar facet joint arthropathy after intra-articular injection of hylan G-F 20. *PM R.* 2009;1(10):908-915.
23. Dimeo F, Bauer M, Varahram I, Proest G, Halter U. Benefits from aerobic exercise in patients with major depression: A pilot study. *British Journal of Sports Medicine.* 2001;35(2):114-117.
24. Dolbow DR, Gorgey AS, Ketchum JM, Moore JR, Hackett LA, Gater DR. Exercise adherence during home-based functional electrical stimulation cycling by individuals with spinal cord injury. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2012;91(11):922-930.

25. Figoni SF, Kunkel CF, Scremin AM, et al. Effects of exercise training on calf tissue oxygenation in men with intermittent claudication. *PM R.* 2009;1(10):932-940.
26. Fraix M. Osteopathic manipulative treatment and vertigo: a pilot study. *PM R.* 2010;2(7):612-618.
27. Francisco G, Chae J, Chawla H, et al. Electromyogram-triggered neuromuscular stimulation for improving the arm function of acute stroke survivors: A randomized pilot study. *Archives of Physical Medicine and Rehabilitation.* 1998;79(5):570-575.
28. Francisco GE, Yozbatiran N, Berliner J, et al. Robot-Assisted Training of Arm and Hand Movement Shows Functional Improvements for Incomplete Cervical Spinal Cord Injury. *American journal of physical medicine & rehabilitation.* 2017;96(10):S171-S177.
29. Fritz SL, Merlo-Rains AM, Rivers ED, et al. An intensive intervention for improving gait, balance, and mobility in individuals with chronic incomplete spinal cord injury: A pilot study of activity tolerance and benefits. *Archives of Physical Medicine and Rehabilitation.* 2011;92(11):1776-1784.
30. Frohm A, Saartok T, Halvorsen K, Renstrom P. Eccentric treatment for patellar tendinopathy: a prospective randomised short-term pilot study of two rehabilitation protocols. *British journal of sports medicine.* 2007;41(7):e7.
31. Gandolfi M, Geroin C, Vale N, et al. Does myofascial and trigger point treatment reduce pain and analgesic intake in patients undergoing onabotulinumtoxinA injection due to chronic intractable migraine? *European journal of physical and rehabilitation medicine.* 2018;54(1):1-12.
32. Gialanella B, Comini L, Gaiani M, Olivares A, Scalvini S. Conservative treatment of rotator cuff tear in older patients: a role for the cycloergometer? A randomized study. *European journal of physical and rehabilitation medicine.* 2018;54(6):900-910.
33. Giusti A, Barone A, Oliveri M, et al. An Analysis of the Feasibility of Home Rehabilitation Among Elderly People With Proximal Femoral Fractures. *Archives of Physical Medicine and Rehabilitation.* 2006;87(6):826-831.
34. Goffredo M, Guanziroli E, Pournajaf S, et al. Overground wearable powered exoskeleton for gait training in subacute stroke subjects: clinical and gait assessments. *European journal of physical and rehabilitation medicine.* 2019;55(6):710-721.
35. Gollie JM, Guccione AA, Panza GS, Jo PY, Herrick JE. Effects of Overground Locomotor Training on Walking Performance in Chronic Cervical Motor Incomplete Spinal Cord Injury: A Pilot Study. *Arch Phys Med Rehabil.* 2017;98(6):1119-1125.
36. Greiner S, Ide J, Van Noort A, et al. Local rhBMP-12 on an Absorbable Collagen Sponge as an Adjuvant Therapy for Rotator Cuff Repair - A Phase 1, Randomized, Standard of Care Control, Multicenter Study: Safety and Feasibility. *Am J Sports Med.* 2015;43(8):1994-2004.

37. Guo YH, Kuan TS, Chen KL, et al. Comparison Between Steroid and 2 Different Sites of Botulinum Toxin Injection in the Treatment of Lateral Epicondylalgia: A Randomized, Double-Blind, Active Drug-Controlled Pilot Study. *Arch Phys Med Rehabil.* 2017;98(1):36-42.
38. Herman T, Giladi N, Gruendlinger L, Hausdorff JM. Six Weeks of Intensive Treadmill Training Improves Gait and Quality of Life in Patients With Parkinson's Disease: A Pilot Study. *Archives of Physical Medicine and Rehabilitation.* 2007;88(9):1154-1158.
39. Huang YC, Chen PC, Tso HH, Yang YC, Ho TL, Leong CP. Effects of kinesio taping on hemiplegic hand in patients with upper limb post-stroke spasticity: a randomized controlled pilot study. *European journal of physical and rehabilitation medicine.* 2019;55(5):551-557.
40. Huber W, Willmes K, Poeck K, Van Vleymen B, Deberdt W. Piracetam as an adjuvant to language therapy for aphasia: A randomized double-blind placebo-controlled pilot study. *Archives of Physical Medicine and Rehabilitation.* 1997;78(3):245-250.
41. Hung JW, Chou CX, Chang HF, et al. Cognitive effects of weight-shifting controlled exergames in patients with chronic stroke: a pilot randomized comparison trial. *European journal of physical and rehabilitation medicine.* 2017;53(5):694-702.
42. Hung JW, Yu MY, Chang KC, Lee HC, Hsieh YW, Chen PC. Feasibility of Using Tetrax Biofeedback Video Games for Balance Training in Patients With Chronic Hemiplegic Stroke. *PM and R.* 2016;8(10):962-970.
43. Iversen MD, Fossel AH, Katz JN. Enhancing function in older adults with chronic low back pain: A pilot study of endurance training. *Archives of Physical Medicine and Rehabilitation.* 2003;84(9):1324-1331.
44. Iwatsu K, Yamada S, Iida Y, et al. Feasibility of neuromuscular electrical stimulation immediately after cardiovascular surgery. *Arch Phys Med Rehabil.* 2015;96(1):63-68.
45. Jerilyn TX, Menz HB, Murley GS, Munteanu SE. Effectiveness of Shoe Stiffening Inserts for First Metatarsophalangeal Joint Osteoarthritis: A Proof-of-Concept Study. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2016;95(2):103-111.
46. Johanson ME, Dairaghi CA, Hentz VR. Evaluation of a Task-Based Intervention After Tendon Transfer to Restore Lateral Pinch. *Arch Phys Med Rehabil.* 2016;97(6 Suppl):S144-153.
47. Kakuda W, Abo M, Momosaki R, et al. Combined therapeutic application of botulinum toxin type A. low-frequency rTMS,andIntensive occupational therapy for post-stroke spastic upper limb hemiparesis. *European Journal of Physical and Rehabilitation Medicine.* 2012;48(1):47-55.
48. Kamm CP, Schmid JP, Muri RM, Mattle HP, Eser P, Saner H. Interdisciplinary cardiovascular and neurologic outpatient rehabilitation in patients surviving transient ischemic attack or stroke with minor or no residual deficits. *Archives of Physical Medicine and Rehabilitation.* 2014;95(4):656-662.

49. Knutson JS, Hansen K, Nagy J, et al. Contralaterally controlled neuromuscular electrical stimulation for recovery of ankle dorsiflexion: a pilot randomized controlled trial in patients with chronic post-stroke hemiplegia. *Am J Phys Med Rehabil.* 2013;92(8):656-665.
50. Kon E, Engebretsen L, Verdonk P, Nehrer S, Filardo G. Clinical Outcomes of Knee Osteoarthritis Treated With an Autologous Protein Solution Injection: A 1-Year Pilot Double-Blinded Randomized Controlled Trial. *Am J Sports Med.* 2018;46(1):171-180.
51. Kumar D, McDermott K, Feng H, et al. Effects of Form-Focused Training on Running Biomechanics: A Pilot Randomized Trial in Untrained Individuals. *PM and R.* 2015;7(8):814-822.
52. Lee M, Lee SH, Kim T, et al. Feasibility of a Smartphone-Based Exercise Program for Office Workers With Neck Pain: An Individualized Approach Using a Self-Classification Algorithm. *Arch Phys Med Rehabil.* 2017;98(1):80-87.
53. Lee SH, Jung HY, Yun SJ, Oh BM, Seo HG. Upper Extremity Rehabilitation Using Fully Immersive Virtual Reality Games With a Head Mount Display: A Feasibility Study. *PM R.* 2020;12(3):257-262.
54. Lehmann I, Thaler I, Luder G, et al. Standardized, comprehensive, hospital-based circuit training in people with multiple sclerosis: results on feasibility, adherence and satisfaction of the training intervention. *European journal of physical and rehabilitation medicine.* 2020;56(3):279-285.
55. Litterini AJ, Fieler VK, Cavanaugh JT, Lee JQ. Differential effects of cardiovascular and resistance exercise on functional mobility in individuals with advanced cancer: A randomized trial. *Archives of Physical Medicine and Rehabilitation.* 2013;94(12):2329-2335.
56. Liu X, Miller YD, Burton NW, Brown WJ. A preliminary study of the effects of Tai Chi and Qigong medical exercise on indicators of metabolic syndrome, glycaemic control, health-related quality of life, and psychological health in adults with elevated blood glucose. *Br J Sports Med.* 2010;44(10):704-709.
57. Lopes LCC, Mota JF, Prestes J, et al. Intradialytic Resistance Training Improves Functional Capacity and Lean Mass Gain in Individuals on Hemodialysis: A Randomized Pilot Trial. *Arch Phys Med Rehabil.* 2019;100(11):2151-2158.
58. Loyola-Sanchez A, Richardson J, Beattie KA, Otero-Fuentes C, Adachi JD, MacIntyre NJ. Effect of low-intensity pulsed ultrasound on the cartilage repair in people with mild to moderate knee osteoarthritis: A double-blinded, randomized, placebo-controlled pilot study. *Archives of Physical Medicine and Rehabilitation.* 2012;93(1):35-42.
59. Lynch EA, Hillier SL, Stiller K, Campanella RR, Fisher PH. Sensory Retraining of the Lower Limb After Acute Stroke: A Randomized Controlled Pilot Trial. *Archives of Physical Medicine and Rehabilitation.* 2007;88(9):1101-1107.
60. Manias P, Stasinopoulos D. A controlled clinical pilot trial to study the effectiveness of ice as a supplement to the exercise programme for the management of lateral elbow tendinopathy. *British Journal of Sports Medicine.* 2006;40(1):81-85.

61. McCain KJ, Pollo FE, Baum BS, Coleman SC, Baker S, Smith PS. Locomotor Treadmill Training With Partial Body-Weight Support Before Overground Gait in Adults With Acute Stroke: A Pilot Study. *Archives of Physical Medicine and Rehabilitation.* 2008;89(4):684-691.
62. McMahon MA, Vargus-Adams JN, Michaud LJ, Bean J. Effects of amantadine in children with impaired consciousness caused by acquired brain injury: A pilot study. *American Journal of Physical Medicine and Rehabilitation.* 2009;88(7):525-532.
63. Meilahn JR. Tolerability and effectiveness of a neuroprosthesis for the treatment of footdrop in pediatric patients with hemiparetic cerebral palsy. *PM R.* 2013;5(6):503-509.
64. Mendias CL, Enselman ERS, Olszewski AM, et al. The Use of Recombinant Human Growth Hormone to Protect Against Muscle Weakness in Patients Undergoing Anterior Cruciate Ligament Reconstruction: A Pilot, Randomized Placebo-Controlled Trial. *Am J Sports Med.* 2020;48(8):1916-1928.
65. Morone G, Pisotta I, Pichiorri F, et al. Proof of principle of a brain-computer interface approach to support poststroke arm rehabilitation in hospitalized patients: Design, acceptability, and usability. *Archives of Physical Medicine and Rehabilitation.* 2015(3):S71-S78.
66. Moseley Jr JB, Wray NP, Kuykendall D, Willis K, Landon G. Arthroscopic treatment of osteoarthritis of the knee: A prospective, randomized, placebo-controlled trial: Results of a pilot study. *American Journal of Sports Medicine.* 1996;24(1):28-34.
67. Mountain AD, Kirby RL, Eskes GA, et al. Ability of People With Stroke to Learn Powered Wheelchair Skills: A Pilot Study. *Archives of Physical Medicine and Rehabilitation.* 2010;91(4):596-601.
68. Nguyen S, McKay A, Wong D, et al. Cognitive Behavior Therapy to Treat Sleep Disturbance and Fatigue After Traumatic Brain Injury: A Pilot Randomized Controlled Trial. *Arch Phys Med Rehabil.* 2017;98(8):1508-1517 e1502.
69. O'Connell DG, Barnhart R. Improvement in wheelchair propulsion in pediatric wheelchair users through resistance training: A pilot study. *Archives of Physical Medicine and Rehabilitation.* 1995;76(4):368-372.
70. Ohberg L, Alfredson H. Ultrasound guided sclerosis of neovessels in painful chronic Achilles tendinosis: Pilot study of a new treatment. *British Journal of Sports Medicine.* 2002;36(3):173-175.
71. Oosting E, Jans MP, Dronkers JJ, et al. Preoperative Home-Based Physical Therapy Versus Usual Care to Improve Functional Health of Frail Older Adults Scheduled for Elective Total Hip Arthroplasty: A Pilot Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation.* 2012.
72. Pang MY, Harris JE, Eng JJ. A community-based upper-extremity group exercise program improves motor function and performance of functional activities in chronic stroke: A randomized controlled trial. *Archives of Physical Medicine and Rehabilitation.* 2006;87(1):1-9.

73. Parke S, Hough CL, A EB. The Feasibility and Acceptability of Virtual Therapy Environments for Early ICU Mobilization. *PM and R.* 2020.
74. Protas EJ, Holmes SA, Qureshy H, Johnson A, Lee D, Sherwood AM. Supported treadmill ambulation training after spinal cord injury: A pilot study. *Archives of Physical Medicine and Rehabilitation.* 2001;82(6):825-831.
75. Rabadi MH, Aston CE. Effect of Transcranial Direct Current Stimulation on Severely Affected Arm-Hand Motor Function in Patients After an Acute Ischemic Stroke: A Pilot Randomized Control Trial. *American journal of physical medicine & rehabilitation.* 2017;96(10):S178-S184.
76. Rabago D, Lee KS, Ryan M, et al. Hypertonic dextrose and morrhuate sodium injections (prolotherapy) for lateral epicondylosis (tennis elbow): results of a single-blind, pilot-level, randomized controlled trial. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2013;92(7):587-596.
77. Richards CL, Malouin F, Wood-Dauphinee S, Williams JI, Bouchard JP, Brunet D. Task-specific physical therapy for optimization of gait recovery in acute stroke patients. *Archives of Physical Medicine and Rehabilitation.* 1993;74(6):612-620.
78. Ridgel AL, Peacock CA, Fickes EJ, Kim CH. Active-assisted cycling improves tremor and bradykinesia in Parkinson's disease. *Archives of Physical Medicine and Rehabilitation.* 2012;93(11):2049-2054.
79. Sakakibara BM, Miller WC, Souza M, Nikolova V, Best KL. Wheelchair skills training to improve confidence with using a manual wheelchair among older adults: A pilot study. *Archives of Physical Medicine and Rehabilitation.* 2013;94(6):1031-1037.
80. Sayers SP, Bean J, Cuoco A, LeBrasseur NK, Jette A, Fielding RA. Changes in function and disability after resistance training: Does velocity matter? A pilot study. *American Journal of Physical Medicine and Rehabilitation.* 2003;82(8):605-613.
81. Schenkman ML, Jordan S, Akuthota V, et al. Functional movement training for recurrent low back pain: lessons from a pilot randomized controlled trial. *PM R.* 2009;1(2):137-146.
82. Schnitzer TJ, Kim K, Marks J, Yeasted R, Simonian N, Chen D. Zoledronic Acid Treatment After Acute Spinal Cord Injury: Results of a Randomized, Placebo-Controlled Pilot Trial. *PM R.* 2016;8(9):833-843.
83. Silber M, Chung TS, Varghese G, Hinterbuchner C, Bailey M, Hivry N. Pneumatic orthosis: pilot study. *Archives of Physical Medicine and Rehabilitation.* 1975;56(1):27-32.
84. Slack D, Nelson L, Patterson D, Burns S, Hakimi K, Robinson L. The feasibility of hypnotic analgesia in ameliorating pain and anxiety among adults undergoing needle electromyography. *American Journal of Physical Medicine and Rehabilitation.* 2009;88(1):21-29.
85. Smith PS, Dinse HR, Kalisch T, Johnson M, Walker-Batson D. Effects of Repetitive Electrical Stimulation to Treat Sensory Loss in Persons Poststroke. *Archives of Physical Medicine and Rehabilitation.* 2009;90(12):2108-2111.

86. Sommers J, Klooster E, Zoethout SB, et al. Feasibility of Exercise Testing in Patients Who Are Critically Ill: A Prospective, Observational Multicenter Study. *Arch Phys Med Rehabil.* 2019;100(2):239-246.
87. Sullivan KJ, Knowlton BJ, Dobkin BH. Step training with body weight support: Effect of treadmill speed and practice paradigms on poststroke locomotor recovery. *Archives of Physical Medicine and Rehabilitation.* 2002;83(5):683-691.
88. Szturm T, Imran Z, Pooyania S, Kanitkar A, Mahana B. Evaluation of a Game Based Tele Rehabilitation Platform for In-Home Therapy of Hand-Arm Function Post Stroke: Feasibility Study. *PM and R.* 2020.
89. Takacs J, Krowchuk NM, Garland SJ, Carpenter MG, Hunt MA. Dynamic Balance Training Improves Physical Function in Individuals With Knee Osteoarthritis: A Pilot Randomized Controlled Trial. *Arch Phys Med Rehabil.* 2017;98(8):1586-1593.
90. Treger I, Faran S, Ring H. Robot-assisted therapy for neuromuscular training of sub-acute stroke patients. A feasibility study. *European Journal of Physical and Rehabilitation Medicine.* 2008;44(4):431-435.
91. Trinh L, Wilson R, Williams HM, Sum AJ, Naylor PJ. Physicians promoting physical activity using pedometers and community partnerships: a real world trial. *Br J Sports Med.* 2012;46(4):284-290.
92. van Rooijen SJ, Molenaar CJL, Schep G, et al. Making Patients Fit for Surgery: Introducing a Four Pillar Multimodal Prehabilitation Program in Colorectal Cancer. *American journal of physical medicine & rehabilitation.* 2019;98(10):888-896.
93. Vincent F, Labourey JL, Leobon S, Antonini MT, Lavau-Denes S, Tubiana-Mathieu N. Effects of a home-based walking training program on cardiorespiratory fitness in breast cancer patients receiving adjuvant chemotherapy: A pilot study. *European Journal of Physical and Rehabilitation Medicine.* 2013;49(3):319-329.
94. Walker ML, Ringleb SI, Maihafer GC, et al. Virtual Reality-Enhanced Partial Body Weight-Supported Treadmill Training Poststroke: Feasibility and Effectiveness in 6 Subjects. *Archives of Physical Medicine and Rehabilitation.* 2010;91(1):115-122.
95. Weeks DL, Whitney AA, Tindall AG, Carter GT. Pilot Randomized Trial Comparing Intersession Scheduling of Biofeedback Results to Individuals with Chronic Pain: Influence on Psychologic Function and Pain Intensity. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2015;94(10):869-878.
96. Williams SB, Brand CA, Hill KD, Hunt SB, Moran H. Feasibility and Outcomes of a Home-Based Exercise Program on Improving Balance and Gait Stability in Women With Lower-Limb Osteoarthritis or Rheumatoid Arthritis: A Pilot Study. *Archives of Physical Medicine and Rehabilitation.* 2010;91(1):106-114.
97. Wu M, Kim J, Gaebler-Spira DJ, Schmit BD, Arora P. Robotic Resistance Treadmill Training Improves Locomotor Function in Children With Cerebral

Palsy: A Randomized Controlled Pilot Study. *Archives of Physical Medicine and Rehabilitation.* 2017;98(11):2126-2133.

98. Wu M, Landry JM, Schmit BD, Hornby TG, Yen SC. Robotic resistance treadmill training improves locomotor function in human spinal cord injury: A pilot study. *Archives of Physical Medicine and Rehabilitation.* 2012;93(5):782-789.
99. Yeung A, Lepoutre V, Wayne P, et al. Tai chi treatment for depression in Chinese Americans: a pilot study. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2012;91(10):863-870.
100. Zaenker P, Favret F, Lonsdorfer E, Muff G, de Seze J, Isner-Horobeti ME. High-intensity interval training combined with resistance training improves physiological capacities, strength and quality of life in multiple sclerosis patients: a pilot study. *European journal of physical and rehabilitation medicine.* 2018;54(1):58-67.

SUPPLEMENTAL DIGITAL CONTENT 3

Supplemental Table 1. Overview with illustrative examples of feasibility conclusions as reported by authors of the included studies.

Scenario	Type of conclusion	Examples
Implications of feasibility findings on a main study discussed (n = 57)	**1. Adequate –** feasible as is	• In a 2003 'pilot' study evaluating various feasibility outcomes (e.g., acceptance, adherence, feedback from patients, retention rate, and data for a sample size estimate), authors conclude that the intervention is a "safe and easily accessible method of enhancing general function and mood in this sample of elders." and do not comment on any needed modifications for the design of a larger study. Authors also tabulate sample size calculations for a randomized trial for five functional outcomes on the basis of the findings, "The observational design does permit us to calculate the likely effect sizes (and sample sizes) for an RCT." and "The SF-36 physical functioning scale was the most responsive instrument". [43] • In a 2008 'feasibility' study with findings based on acceptance, adherence and feedback from patients, authors state "that the use of this system was associated with high patient's satisfaction" and that "The Reo™ Therapy System was found to be valued by patients". No modifications are indicated and authors conclude that "Further research is necessary in order to identify the most efficient balance of Reo Therapy and traditional therapy methods" since the intervention was evaluated alongside standard of care rehabilitation sessions. [90] • In a 2002 'pilot' study, authors measured patient satisfaction in addition to effectiveness and conclude that the intervention is "safe and reasonably simple to perform in the radiology room". No modifications to the design are noted, and although there was no prior mention of plans for a future larger study, authors state, albeit vaguely, that "We have started appropriately controlled and blinded randomised studies to evaluate further the effects of this type of treatment." also citing the lack of a control group and blinding as limitations. [70]

	2. Adequate – feasible with modifications	• In 2010 'pilot' study, authors measured acceptance, data completeness, recruitment and retention rates. They state that "the current study would have benefitted from having additional outcome measures, specifically CDP", which is validated. With regard to a larger study, "this study demonstrated that is feasible to recruit a population of patients with vertigo, complete a course of OMT, and collect data concerning its potential efficacy by using the DHI. A similar randomized control trial that uses separate physicians for evaluation and treatment is feasible and warranted to determine the efficacy of OMT in vertigo." [26] • In a 2019 pilot study assessing feasibility (primarily acceptance), authors conclude "The treatment was considered comfortable and enjoyable (mean score >5) and moderately painful and strenuous (mean scores <3). Moreover, the majority of the patients perceived the exoskeleton as useful; they would recommend it; and they would like to do further OEAGT in the near future." although it is unclear if the reported means represent meaningful scores. Authors also state that the "TUG test was too challenging for our subacute stroke patients." with "only 16.67% subjects were able to perform the test", and suggest several alternative measures for larger studies. Authors conclude that the study "demonstrates that future RCT studies with a larger population are recommended" and that the current study is "part of a wider RCT" but do not elaborate if the current study is an internal pilot or vanguard study. [34] • In a 2020 'feasibility' study evaluating acceptance, adherence, patient feedback and safety, authors conclude that the intervention was feasible without serious adverse events and high satisfaction. Authors comment on numerous modifications including a structured questionnaire to track adverse events (which may have been undetected), adjustments to the program (e.g., linking to more hand motion devices with haptics, improving the virtual space to show the whole upper extremity versus only a hand which "can interfere with the patient's embodiment. This may affect the actual therapeutic effect"). Authors also highlight the need to address the generational gap in future investigations

		"relatively young patients showed more interest in VR rehabilitation during the recruitment process...may be considered another issue confronting VR rehabilitation." [53] • In a 2019 study investigating various aspects of feasibility, clear linkage to a future study is provided, including plans to evaluate logistics for an international randomized trial. Authors conclude that the intervention is safe and feasible, noting various adaptations to the current protocol (e.g., timing of dose) as well as assessing the need for the psychologist component with a questionnaire, "on evaluation, there was not much added value because fear for the operation and the cancer diagnosis overall was realistic and weekly counseling by the case manager (nurse specialist in colorectal care) would suffice. We now use validated questionnaires PHQ-9 and GAD-7 for screening to select individual patients for psychologist consultation." Authors also provide trial registration information for the main, international study. [92] • In a 2009 'pilot' study with plans to collect data for a sample size estimate for a definitive study, authors provide the following power calculation "A future study could be designed with a sample size of 12 (6 in each arm) to detect a difference in means of 1.64 on the CRS-R, using a two-group *t* test (crossover analysis of variance) with a 0.05 two-sided significance level and 80% power. A more conservative calculation based on a difference of means of 1.20 would require a total sample size of 20." Authors address issues faced and considerations for future recruitment efforts "most were willing to take part in a crossover design. These families' attitudes support the difficulties in a parallel group design in this setting", also noting the need to account for high dropout rates due to medical complications and desire for open-label use. Due to slow recruitment, authors state there is a "need for a four- to five-site multicenter study" and discuss criticisms of crossovers with agents like amantadine which may result in "irreversible improvement". However, the authors conclude that the findings from the washout period "suggests that a crossover design is likely to detect meaningful changes related to the use of amantadine". [62]

		• In a 2010 'feasibility' study with aims "to facilitate study replication and clinical implementation", authors conclude that "outpatients undergoing radiation treatment can participate in, enjoy, and exercise training." Authors note that combining sessions with radiation center visits was a logistical benefit and offer suggestions for future research, "Minor modifications might achieve greater and more lasting benefit. For example, the study protocol did not provide for individualization of the exercise regimen or performance-based program advancement. Both of these are standard elements of conventional PT and could be reasonably expected to enhance benefit. Furthermore, inclusion of an individualized and structured post-PT home exercise program, another essential component of conventional outpatient PT, may have facilitated continued improvement after completion of the intervention. Aerobic conditioning might be added to the regimen because there is evidence supporting the benefits of this modality for fatigue." Authors also comment on the need to measure functional status objectively in addition to subjective measures, and "to examine the capacity of telephone, internet, or in-person follow-up to extend gains beyond the study interval." [18]
	3. Adequate – implementation-oriented discussion (rather than future research)	• In a 2012 study evaluating feasibility, authors collected feedback from patients, physicians, and community coordinators, stating the intervention was successful and "demonstrated the feasibility of implementing a pedometer intervention in partnership with the community." Authors highlight that the "Long-term maintenance of this intervention approach would require support for ongoing communication between the physicians and CAS and the necessary resources to encourage PA and follow-up with patients" and that "lack of time was viewed by the physicians as the greatest challenging to implementing the PA intervention into their practice". The authors also comment on issues with bias in the self-reported measures used, concluding that the findings "support a broader implementation trial". [91] • In a 2014 study evaluating the feasibility of integrating a program into clinical practice, authors evaluated various aspects including acceptance, practicability,

		and logistics of integrating the program with existing structures and organizational aspects of the center "without significant adaptation and interference". The authors conclude that "The integration of the presented program into an existing cardiovascular prevention center was feasible, safe, and reasonable due to the largely overlapping therapeutic measures on vascular prevention.", and that the program was "well accepted with 90% of the patients completing it. This is a great step toward the creation of a vascular prevention and rehabilitation center offering such services to patients with various vascular diseases.", also noting that cost-effectiveness needs to be determined in future studies. [48] • In a 2020 study evaluating feasibility, authors conclude that the program was feasible and address considerations for implementation, "Travel and time issues as well as Uthoff's phenomenon during summer month were main reasons for dropouts during the intervention period. We therefore recommend pausing the program during the summer". In addition to reporting relevant outcomes (adherence, patient satisfaction and feedback), authors also report the use of the TIDieR guideline to encourage optimal reproducibility of the intervention in different settings, "Current research in rehabilitation often cannot be translated into clinical practice because the applied training programs are not reported adequately and therefore cannot be reproduced. We therefore developed a, within this paper, clearly described and reproducible training program with regard to the design of the program and the individual workstations." [54]
Implications of feasibility findings on a main study not discussed (n = 43)	**4. Inadequate –** comments about future research in absence of a priori feasibility objectives or outcomes	• In a 2017 'pilot' study to determine effectiveness with no a priori feasibility objectives or outcomes, authors provide general recommendations with regard to future research based on the pilot's limitations (i.e., bigger sample, need for a control group and longer study duration), "In the future, a longer duration study with an adequate sample size, including a group with video game training only and using a more standardized neuropsychological assessment may clearly establish the potential cognitive benefits of these interventions." [41]

		• In a 2017 'pilot' study with the objective to determine efficacy, feasibility-related outcomes were first mentioned in the results (e.g., adherence by therapists, and a measure of exercise output—distance walked—which was only clarified as an indicator of compliance in the results). Authors also state that "Power calculations were not performed because the estimates from this pilot study will be used to determine adequate sample size for a larger trial", but further assert that the pilot provides empirical evidence "that CBT is effective in alleviating sleep, daily fatigue levels, and depression in a TBI sample." The statement that "Treatment effect size, the magnitude of change from baseline within groups, was estimated with Hedges *g* to adjust for small sample bias" also suggests that the study may have been a small study later labeled as a pilot, although this is unclear. [68] • In a 2010 'pilot' study with the objective to determine short-term efficacy, some feasibility-related outcomes were only mentioned in the results (e.g., acceptance, feedback by patients), and intentions for the pilot to guide the design of a larger study were first presented in the discussion, "An equivalency study of this type would require large numbers of subjects, and given the costs involved, the authors believe an initial pilot study was justified." Authors highlight that the study did not support the use of the higher dose as well as challenges due to the large number of lost to follow-ups, and with conducting similar larger studies "However, large studies of this type are challenging, particularly if carried out independent of pharmaceutical support, because of cost and the need for FDA clearance." [10]
	5. Inadequate – efficacy/effectiveness-oriented discussion, or discussion of other, non-measured feasibility outcomes	• In a 2020 'pilot' study to determine safety, tolerance, and preliminary efficacy data, authors largely provide a discussion from the perspective of efficacy, "additional trials that consider longer and perhaps higher doses of HGH in a larger population of male and female patients, along with neuromuscular and additional functional evaluations, would help to further determine the therapeutic potential of HGH to improve outcomes for patients who undergo ACLR." Authors report a large effect size based on data from 19 patients, "The

		primary outcome measure for this study was knee extension strength at 6 months. In support of our hypothesis, we observed a 33% increase in absolute torque and a 29% increase in relative torque in the HGH group compared with the placebo group at 6 months after ACLR." Based on this, the safety committee (blinded to the treatment assignments) decided to stop the study, "Enrollment was discontinued before reaching the preplanned number of 17 per group, which can result in overestimates of the effect in decisive trials, but we thought that this was the correct decision, as this was a pilot trial." [64] • In a 2014 'pilot' study to determine safety and efficacy, authors comment on additional, potentially measurable outcomes (i.e., which could have been captured with more preplanning) stating "Some feelings (eg, scalp sensations) may have differed in the placebo and real conditions." and so "the future use of a questionnaire for study participants and evaluating physicians is recommended to help verify that blinding is maintained throughout the conduct of the study." [19]

Note: numbered references appear in SUPPLEMENTAL DIGITAL CONTENT 2

CHAPTER 3: A pilot study of methodological studies

3.0 Preface

The following pages include the published study covering Chapter 3. Reprinted in accordance with the open access copyright license and permission from the author.

Citation:

Lawson, D. O., Leenus, A., & Mbuagbaw, L. (2020). Mapping the nomenclature, methodology, and reporting of studies that review methods: a pilot methodological review. Pilot and Feasibility Studies, 6, 13. doi:10.1186/s40814-019-0544-0

Lawson *et al. Pilot and Feasibility Studies* (2020) 6:13
https://doi.org/10.1186/s40814-019-0544-0

Pilot and Feasibility Studies

RESEARCH Open Access

Mapping the nomenclature, methodology, and reporting of studies that review methods: a pilot methodological review

Daeria O. Lawson[1*], Alvin Leenus[2] and Lawrence Mbuagbaw[1,3]

Abstract

Background: A relatively novel method of appraisal, methodological reviews (MRs) are used to synthesize information on the methods used in health research. There are currently no guidelines available to inform the reporting of MRs.

Objectives: This pilot review aimed to determine the feasibility of a full review and the need for reporting guidance for methodological reviews.

Methods: Search strategy: We conducted a search of PubMed, restricted to 2017 to include the most recently published studies, using different search terms often used to describe methodological reviews: "literature survey" OR "meta-epidemiologic* review" OR "meta-epidemiologic* survey" OR "methodologic* review" OR "methodologic* survey" OR "systematic survey."
Data extraction: Study characteristics including country, nomenclature, number of included studies, search strategy, a priori protocol use, and sampling methods were extracted in duplicate and summarized.
Outcomes: Primary feasibility outcomes were the sensitivity and specificity of the search terms (criteria for success of feasibility set at sensitivity and specificity of ≥ 70%).
Analysis: The estimates are reported as a point estimate (95% confidence interval).

Results: Two hundred thirty-six articles were retrieved and 31 were included in the final analysis. The most accurate search term was "meta-epidemiological" (sensitivity [Sn] 48.39; 95% CI 31.97–65.16; specificity [Sp] 97.56; 94.42–98.95). The majority of studies were published by authors from Canada (n = 12, 38.7%), and Japan and USA (n = 4, 12.9% each). The median (interquartile range [IQR]) number of included studies in the MRs was 77 (13–1127). Reporting of a search strategy was done in most studies (n = 23, 74.2%). The use of a pre-published protocol (n = 7, 22.6%) or a justifiable sampling method (n = 5, 16.1%) occurred rarely.

Conclusions: Using the MR nomenclature identified, it is feasible to build a comprehensive search strategy and conduct a full review. Given the variation in reporting practices and nomenclature attributed to MRs, there is a need for guidance on standardized and transparent reporting of MRs. Future guideline development would likely include stakeholders from Canada, USA, and Japan.

Keywords: Feasibility, Guidelines, Methodological review, Nomenclature, Pilot, Reporting

* Correspondence: lawsod3@mcmaster.ca
[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
Full list of author information is available at the end of the article

Background

Health researchers, methodologists, and policymakers rely on primary studies or evidence syntheses (e.g., systematic reviews) to provide summaries of evidence for decision-making [1, 2]. However, the credibility of this evidence depends on how the studies were conducted and reported. Therefore, critical appraisal of health research methodology is an important tool for researchers and end-users of evidence. As such, certain studies exist solely to help synthesize methodological data about the design, analysis, and reporting of primary and secondary research. These studies can be referred to, for the purposes of this paper, as methodological reviews (MRs) and represent an efficient way of assessing research methods and summarizing methodological issues in the conduct, analysis, and reporting of health research. Collating primary and secondary research in this way can help to identify reporting and methodological gaps, generate empirical evidence on the state of or quality of conduct and reporting, and inform the development of reporting and methodological standards. MRs are highly informative because they allow researchers to evaluate study methods; assess adherence, quality, and completeness of reporting (e.g., reporting adherence to Consolidated Standards of Reporting Trials, CONSORT); document and assess the variety of methods used or approaches to analyses (e.g., statistical approaches for handling missing data in cluster randomized trials); demonstrate changes in reporting over time (e.g., since the introduction of a specific guideline); demonstrate consistency between study abstracts/trial registries and their full texts; and many other issues [3–8]. In this way, MRs are indispensable to high-quality health research by allowing researchers to identify inappropriate research method practices and propose solutions.

Reporting guidelines are important tools in improving the reporting and conduct of health research, and many exist for various study designs. Currently, the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) network is the leading authority in reporting guidelines for health research [9]. As of September 2019, this website lists 418 guidelines, with another 74 currently under development. There is empirical evidence that publication of reporting guidelines improves reporting, but this is often contingent on journal endorsement as well as the period of time since publication [10–12]. However, there is no published guidance for reporting methodological reviews despite an increase in their development and usage [13].

Murad and Wang have proposed a checklist which is an adaptation of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), a widely endorsed guideline developed to help report systematic reviews [14, 15]. While its use would ensure that a standardized and transparent approach in reporting is followed, this does not address the various typologies of MRs including the variety of approaches used to conduct these studies. For example, MRs that use a before-after design, interrupted time series, or random sampling approaches would be a poor fit for this tool [7, 16–18]. Likewise, studies in which the unit of analysis is not the "study" would require more specific guidance (e.g., some MRs investigate multiple subgroup analyses within the same study) [18, 19]. Further, some MRs report formal sample size estimations, and it is unclear whether this should be recommended for all MRs. In line with this thinking, recent correspondences with the Journal of Clinical Epidemiology highlight that methodological studies cannot always be classified as systematic reviews, but instead represent their own branch of evidence synthesis methodology requiring specially tailored reporting guidance [20].

To facilitate the development of reporting guidelines for health research, Moher et al. propose a five-phase, 18-step strategy [10]. One important step in this process is a review of the literature, including seeking evidence on the quality of published research articles and identifying information related to sources of bias in reporting. However, some concerns exist with reviewing the literature. First, the literature on methodological reviews is elusive and can be found in any journal or database. Second, given the relative novelty and rapid development of the field, there is no formally accepted nomenclature to guide a literature search. We therefore deemed it necessary to conduct a pilot methodological review of methodological reviews to inform the feasibility of a broader, full methodological review based on our ability to

1. Determine the appropriate nomenclature for identification of methodological reviews
2. Determine a preliminary need for guidance, based on inconsistencies in reporting

Pilot studies of research syntheses, contrary to pilot studies of classical "primary" research, aim to: establish the need for a full review, establish the value of the methods used, and to identify, clarify, and review any problems with the processes and instruments. They can also be used to identify conceptual, methodological, and practical problems that need to be addressed in a full review. In this way, piloting research syntheses maximizes validity and efficiency [21, 22]. The research questions we sought to answer in this pilot review, where they fit in the larger scheme of the project objectives, and their implications for a full review (i.e., larger study to further explore the observed methodological variations in a broader sample of MRs) are outlined in Table 1. The listed research questions were evaluated in and applied to the MRs included in this pilot.

Methods

Study design

The methods reported are in line with current guidance on piloting evidence syntheses, by way of a "mini-review" all the way through [21]. We conducted a pilot

Table 1 Pilot research questions and implications for a full review

Pilot review objectives	Research questions	Implications for feasibility of full review	Metrics/threshold
Determine the appropriate nomenclature for accurate identification of methodological reviews	Which search terms yield methodological reviews?	Identifying a list of terms that yields methodological reviews will inform the search strategy in the full review	Sensitivity/specificity ≥ 70%
Determine the need for methodological review reporting guidelines	Are research methods specified a priori?	Inconsistent pre-specification of methods would indicate the need for a full review	≤ 70% with published protocols
	How many databases are searched?	Wide variation in the numbers of databases searched would indicate the need for a full review	Coefficient of variation ~ 1 (i.e., spread in results relative to the mean)
	Are search time limits justified?	Inappropriate justification of search time limits would imply the need for a full review	≤ 70% justify search limits
	Is the sample size justified?	Inappropriate justification of sample size for MRs designed as analytical studies (e.g., before-after comparisons, regression-based analyses) would imply the need for a full review	≤ 70% justify sample size or perform sample size calculation
	Is a formal sample size calculation performed?	Inappropriate justification of sample size for MRs designed as analytical studies (e.g., before-after comparisons, regression-based analyses) would imply the need for a full review	
	Is a random sample of studies used?	Use of different sampling approaches to select a subset of studies from a larger group would indicate the need for a full review	Among studies where the goal was not to capture all available studies, ≤ 70% use a random sampling approach
	Do research methods or authors suggest generalizable findings?	Lack of clear approaches to reporting generalizability would indicate the need for a full review	≤ 70% discuss the generalizability of findings

methodological review of a sample of methodological reviews published in 2017.

Eligibility criteria

We included articles that fulfilled all of the following criteria:

- Are in the domain of clinical research with human participants
- Could be classified as secondary research (i.e., investigating other studies/primary research)
- Investigate methods or reporting issues

Search strategy

We conducted a search of PubMed—a public search engine which retrieves medical literature from the MEDLINE database—from January 1, 2017 to December 31, 2017 (i.e., the most recent complete year) using terms often used to refer to MRs: "literature survey" OR "meta-epidemiologic* review" OR "meta-epidemiologic* survey" OR "methodologic* review" OR "methodologic* survey" OR "systematic survey". To maintain a focused search as was intended for the scope of this pilot, phrase searching was used to restrict the volume of hits with wildcard searching.

Study selection

One reviewer (DOL) screened the titles and abstracts of retrieved articles in a reference manager program (EndNote X7.8, Philadelphia: Clarivate Analytics; 2016) for the type of study and whether any nomenclature was present in either or both the title and abstract. Studies identified as methodological reviews were screened in duplicate (DOL and AL) and verified for eligibility in the full texts using a spreadsheet (Microsoft Excel for Mac v.16.15, Washington: Microsoft Corporation; 2018).

Data extraction

Two reviewers (DOL and AL) extracted data from MRs in a spreadsheet in duplicate including the first author name, country of primary affiliation (for > 1 co-first authors, the affiliation of the first listed author was taken; for > 1 affiliations of the first author, the senior author's affiliation was taken), nomenclature in the title, nomenclature in the abstract, nomenclature in the methods section, total number of included records (e.g., abstracts, instruments, journals, meta-analyses, reviews, and trials), if the sample size was calculated, if the authors referenced a published protocol, the databases searched for record inclusion, if there was justification of search time limits, whether the search strategy was reported (or referenced elsewhere), and if sampling of records was random.

Given our concerns with sampling issues, we sought additional information on generalizability. Generalizability was guided by the response to the following question: "Do these findings represent the total population of studies that the sample was drawn from?". We classified studies as likely generalizable if they met some or all of the following criteria:

- Used multiple databases

- Justified their sample size (e.g., provide details for a sample size calculation)
- Selected a random sample of records (where applicable)

We classified studies as unlikely to be generalizable if they

- Used only one database
- Used selected journals (e.g., only high impact)
- Used very stringent eligibility criteria

Disagreements were resolved through discussion, and if reviewers could not come to an agreement on conflicts, a third reviewer (LM) adjudicated as necessary.

Inter-rater agreement

We assessed the level of agreement between reviewers using Cohen's kappa (κ) for inter-rater reliability for two raters. Agreement was calculated for yes/no and numerical fields at the full-text screen and data extraction levels. The index value was interpreted as no agreement (0–0.20), minimal agreement (0.21–0. 39), weak agreement (0.40–0.59), moderate agreement (0.60–0.79), strong agreement (0.80–0.90), and almost perfect agreement (above 0.90) [23].

Data analysis

We summarized and reported the descriptive statistics, including frequencies and percentages for count and categorical variables, and means with standard deviation (SD) for continuous variables.

In order to determine the best search strategy, we computed the sensitivity and specificity of each search term to determine which would have the best accuracy in our database of studies identified from our search. We computed the proportion of studies identified with each search term that were actually MRs (i.e., included studies and true positives) and the proportion identified that were not MRs (i.e., excluded studies and false positives). We also computed the proportion of studies not captured by the search term that were not MRs (i.e., excluded studies and true negatives) and the proportion of MRs that were not captured by the search term (i.e., included studies and false negatives). We pooled the sensitivity and specificity estimates for multiple search terms using the parallel testing approach in order to achieve an optimal combination of search terms with a good balance of sensitivity and specificity [24]. Statistical analysis was conducted in IBM SPSS Statistics (IBM SPSS Statistics v.24.0, Armonk: IBM Corporation; 2016) and inter-rater agreement (κ with 95% confidence interval, CI) was calculated using WinPepi [25]. We built a word cloud in WordArt to visualize common terms used for indexing MRs in PubMed [26].

Ethics review

Ethics committee approval and consent to participate was not required as this study used previously published non-human data.

Results

There were 236 articles retrieved from the PubMed search of which 31 were included in the final quantitative and qualitative analysis (see Fig. 1 for study flow diagram with reasons for exclusion) [27–57]. There was moderate inter-rater agreement between reviewers before consensus (κ = 0.78; 95% CI 0.73–0.82).

The study characteristics of all included studies are outlined in the Appendix.

Characteristics of included methodological reviews

Many of the authors were from Canadian institutions (n = 12, 38.7%), followed by Japan and USA (n = 4, 12.9% each). Based on the previously defined criteria, we scored ten studies (32.3%) as generalizable [28, 30, 38, 40, 49, 51–54, 57]. Only three studies (9.7%, two of which we scored as generalizable) commented on generalizability and reported their own work as generalizable, either to the subject area (e.g., venous ulcer disease), to a clinical area, or in general terms [27, 30, 38].

Nomenclature

Of the 31 included studies, 77.4% (n = 24) presented the study nomenclature in their title. The terms found in the titles and abstracts of all retrieved articles that were used to describe the study are represented in Fig. 2. The most accurate search terms were "meta-epidemiological" (sensitivity [Sn] 48.39; 95% CI 31.97–65.16; specificity [Sp] 97.56; 94.42–98.95), "systematic survey" (Sn 45.16; 95% CI 29.16–62.23; Sp 76.10; 95% CI 69.81–81.42) "systematic review" (Sn 12.90; 95% CI 5.13–28.85; Sp 93.17; 95% CI 88.86–95.89), and "literature survey" (Sn 6.45; 95 CI 1.79–20.72; Sp 33.66; 95% CI 27.54–40.38), among 12 different types or combinations of nomenclature cited (Fig. 2). The combined sensitivity and specificity for the six terms attributed to MRs was 100% and 0.99%, respectively.

The words "survey" and "systematic" (n = 18 each), "meta-epidemiologic*" (n = 15), and "review" and "study" (n = 7 each) were the most frequent in the word cloud (Appendix). Five (16.1%) of these studies used more than one name to describe their study type in the title or abstract (two names, n = 4, and three names, n = 1).

Methodological features

The mean (standard deviation [SD]) number of databases searched was 2 (1.6) with a minimum of 1 and a maximum of 8 databases. Overall, less than a quarter of studies (n = 7, 22.6%) made a reference to or cited a protocol for the study. Most studies reported the search strategy (n

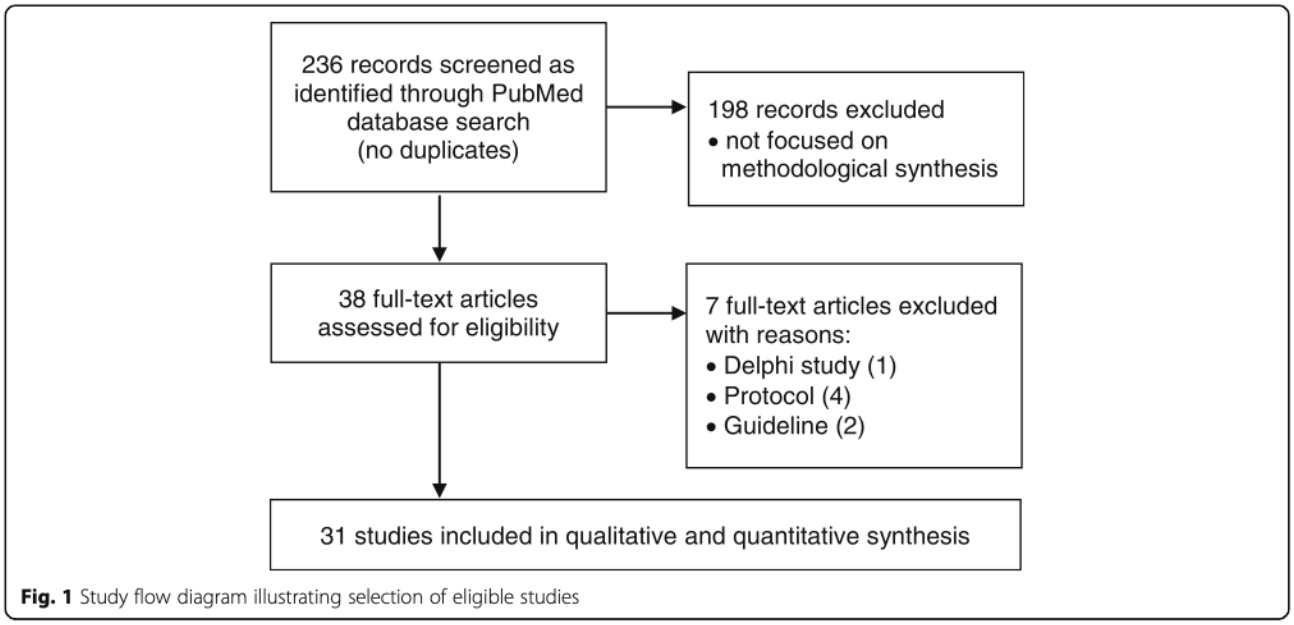


Fig. 1 Study flow diagram illustrating selection of eligible studies

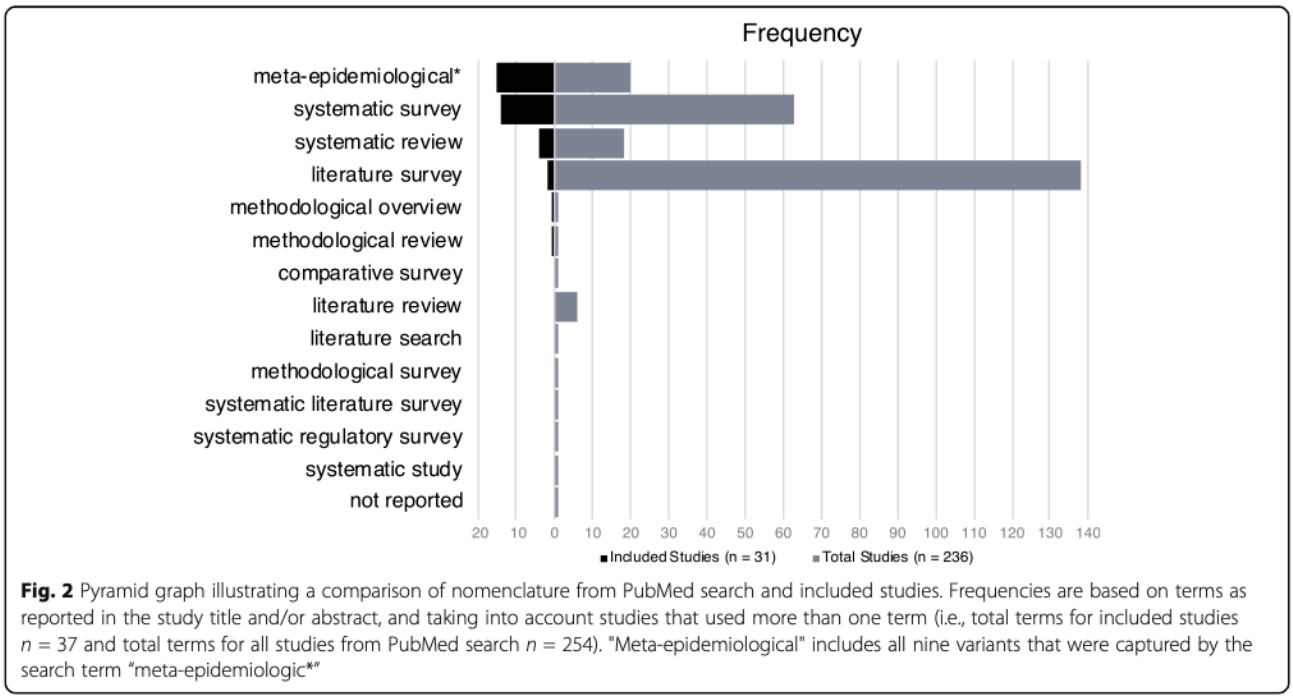


Fig. 2 Pyramid graph illustrating a comparison of nomenclature from PubMed search and included studies. Frequencies are based on terms as reported in the study title and/or abstract, and taking into account studies that used more than one term (i.e., total terms for included studies $n = 37$ and total terms for all studies from PubMed search $n = 254$). "Meta-epidemiological" includes all nine variants that were captured by the search term "meta-epidemiologic*"

= 23, 74.2%) and only eight (25.8%) justified the time limits that were set for the search (Table 2).

Five studies (16.1%) performed an a priori sample size calculation of records to be included, using a variety of methods. Abbade et al. used an estimate for the primary objective from a prior similar study [27], and similarly, Riado Minguez et al. used a prior similar study to determine a target sample size coupled with a power calculation [43]. El Dib et al. sampled enough studies to achieve a CI of ± 0.10 around all proportions [33], and Zhang et al. used a precise CI of ± 0.05 [57]. Kosa et al. incorporated an approach optimized for logistic regression, based on estimates for correlation between covariates [37]. Among the studies that did not aim to summarize data from all available records retrieved in their search (n = 5), all studies (100%) incorporated some randomization strategy to sample the records to be included in their final synthesis [29, 33, 34, 37, 57].

Discussion

In this pilot methodological review, we have established the need for a full review and determined some of the methodological features worth investigating to facilitate the development of a reporting guideline for MRs. Unquestionably, it is highly likely that our search strategy missed MRs characterized by different nomenclature. However, as a result of this pilot we have been able to identify some of the most appropriate search terms to incorporate into a search strategy in the full review. The criteria for success of feasibility and the respective results are outlined in Table 3. The position of this pilot in the larger picture of the development of the reporting guideline is outlined in Fig. 3. Additionally, the forthcoming guideline as a result of this work, the METhodological Review reportIng Checklist (METRIC), has been registered as currently under development with EQUATOR.

The disparity in nomenclature, methods, and reporting in this assessment of 1-year worth of data suggests that a full review is required to better provide a more complete picture of the existing concerns with reporting quality. A number of key features stand out and an appraisal of these concerns would also help to inform the development of guidance. First, there is a growing body of literature pertaining to reviews of methods. A search of PubMed with the term "methodological review" shows that there has been a steady increase in studies indexed as MRs over the past 10 years, with ten in 2007 and 39 in 2017. The increasing number of publications addressing methodological issues in primary and secondary research would suggest that there is interest in understanding and optimizing health research methods. Therefore, the development of consensus-based guidelines in this field is warranted.

Table 2 Methodological features of included methodological reviews (N = 31)

Variable	n (%)
Reported study type (nomenclature) in the "Methods" section	12 (38.7)
Number of databases searched (mean, SD)	2 (1.6)
Number of records included (median, IQR)	77 (13 – 1127)
Reported (or referenced) a protocol	7 (22.6)
Reported (or referenced) a search strategy	23 (74.2)
Justified search time limits	8 (25.8)
Performed a sample size calculation a priori	5 (16.1)
Randomly sampled included records (of n = 5)	5 (100)

IQR interquartile range, *SD* standard deviation

Second, there are inconsistencies in the nomenclature used to describe MRs. In the variety of names currently being attributed to MRs, nomenclature is an issue that must be addressed. This is especially true with the use of labels such as "systematic review"—which is attributed to a specific, well-defined form of evidence synthesis for healthcare studies—or in the case of some studies which have used the term "methodologic," which is not otherwise defined in English dictionaries and which could compromise their detection in searches [58].

Third, there are no methodological standards specific to MRs. Regarding selection bias, it is unclear what processes researchers employ in defining the appropriate eligibility criteria, or in the selection of databases and time periods for screening relevant literature. Readers would be interested in knowing why and how the choice of studies is ideal to answer the research question and the rationale behind such choices should be explicit. Likewise, the approaches to the sampling should be explicit, especially in methodological reviews that do not adopt systematic searches to identify and capture all of the relevant articles. For example, some methodological studies might approach a research question with a before-after, cross-sectional, or longitudinal design to name a few. As a result, it may not be appropriate or necessary to use all of the available studies in these scenarios and a sample of studies may suffice [59, 60].

On a conceptual level, we sought to develop a definition of study generalizability, but this was challenging to operationalize. The aim of this exercise was to help define the scope of inferences that can be made from the findings in MRs. As this area currently lacks specific guidance, we recognize that the appropriateness of the selected criteria (and its applicability to each study) will see ongoing development in subsequent investigations and may be applied differently. Generalizability is strongly tied to the target population and this was not often explicit in MRs, making any inferences challenging. Our approach to defining generalizability could be refined with insights from authors and users of MRs and will vary based on the study in question. There are several ways of addressing this outcome: do the authors identify their study as generalizable? Is the "study topic" (i.e., methodological issue) generalizable to other fields? Are the "results" generalizable to other studies in different fields investigating the same methodological issue? How are these results applicable?

Conversely, we can also consider whether the sample size and number of databases searched are surrogate indicators of

Table 3 Feasibility results for this pilot review

Measure	Target	Observed	Description
Sensitivity/specificity[b]	≥ 70%	Sensitivity, 100% Specificity, 0.99%	Six terms combined gave good sensitivity but compromised specificity
Published protocols[a]	≤ 70%	22.6%	Few studies had pre-specified methods
Coefficient of variation[c]	~ 1	0.8	Fairly consistent number of databases searched
Justification of search limits[a]	≤ 70%	25.8%	Few studies justified their search limits
Justification of sample size or perform sample size calculation[a]	≤ 70%	16.1%	Few studies justified their sample sizes or performed calculations
Use a random sampling approach[c]	≤ 70%	100%	All studies adopted a random sampling approach
Discuss the generalizability of findings[a]	≤ 70%	9.7%	Few studies described how generalizable their findings were

[a]Feasibility criteria met
[b]Feasibility criteria partially met
[c]Feasibility criteria not met

generalizability. These factors could be used to extrapolate to the generalizability of an MR, as is done with clinical trials and systematic reviews, and as we have done in the present study. These questions and the importance of this variable might be answered with a deeper investigation of MRs, as well as feedback and engagement from expert users as we work to develop guidance on reporting. Through expert consensus, and recognizing typological differences, we also plan to optimize the proposed guidance for specific types of MRs (e.g., MRs assessing methods of randomized control trials or systematic reviews). We hope that these approaches will also help to tease out the appropriate definition of generalizability in each case.

Conclusions

We now have a clearer understanding of the terms used to describe methodological reviews and some of the issues that warrant a deeper investigation. In this pilot review, we have highlighted the need for a full review on this topic in order to inform future guidance for reporting methodological reviews. A full review using some of the search terms identified here is feasible. These findings will be used to develop a protocol, which will encompass more databases and years, in order to gain a clearer sense of the landscape of MRs.

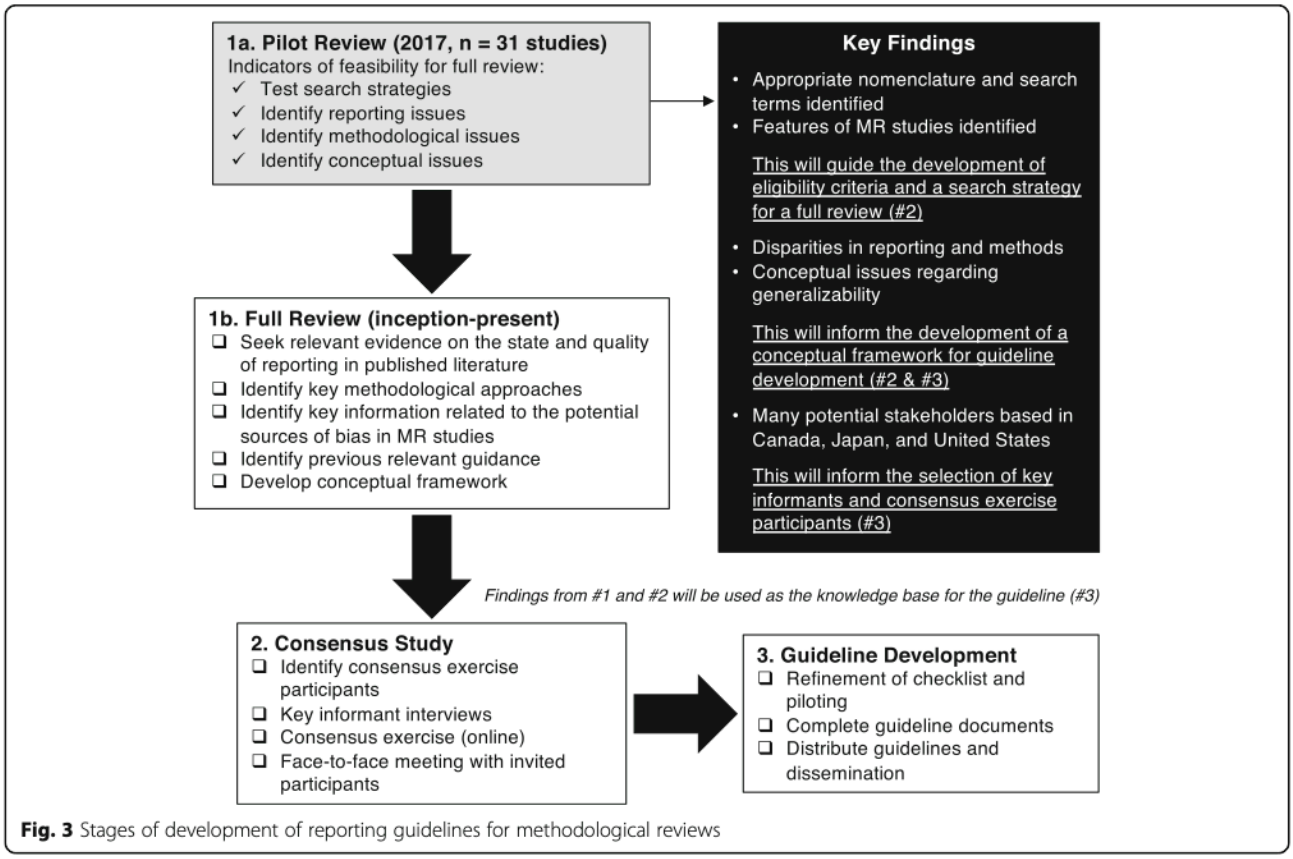


Fig. 3 Stages of development of reporting guidelines for methodological reviews

Appendix

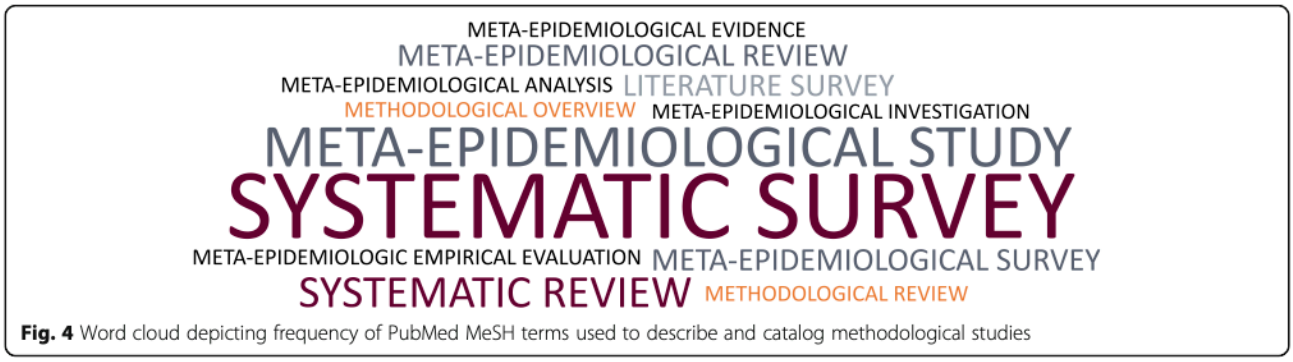


Fig. 4 Word cloud depicting frequency of PubMed MeSH terms used to describe and catalog methodological studies

Table 4 Main characteristics of included methodological reviews (n = 31)

Study	Country	Nomenclature (T/A)	Nomenclature (M)	Databases searched (#)	Records included (#)	Reference to a protocol (Y/N)	Search strategy reported (Y/N)	Justification of search time limits (Y/N)	Sample size calculation (Y/N)	Random sampling (Y/N)
Abbade et al. [27]	Canada	SSu	SSu	1	85	Y	Y	Y	Y (61)	N
Abdul-Khalek et al. [28]	Lebanon	SSu	None	2	57	N	Y	N	N	N
Armijo-Olivo et al. [29]	Canada	MES	MES	1	393	Y	Y	N	N	Y
Bolvig et al. [30]	Denmark	MES	None	1	126	Y	Y	N	N	N
Chase Kruse and Matt Vassar [31]	USA	SSu	SSu	2	35	N	N	N	N	N
Ebrahim et al. [32]	Canada	SR/SSu	None	3	28	Y	Y	N	N	N
El Dib et al. [33]	Canada	SSu	None	1	103	N	Y	N	Y (100)	Y
Ge et al. [34]	China	MES	MESc	1	150	N	Y	N	N	Y
Gorne and Diaz [35]	Argentina	SR/SSu	LSu	1	128	N	Y	N	N	N
Khan et al. [36]	Canada	SSu	SSu	3	48	N	Y	N	N	N
Kosa et al. [37]	Canada	MR, SSu	MR	1	200	N	Y	Y	Y (152)	Y
Kovic et al. [38]	Canada	MESu	MES, SR	2	77	N	Y	Y	N	N
Kuriyama et al. [39]	Japan	LSu	None	1	353	N	N	Y	N	N
Manja et al. [40]	Canada	SSu	None	5	43	N	Y	N	N	N
Papageorgiou et al. [41]	Switzerland	MEE, MO, SR	SR	8	34	Y	Y	N	N	N
Ratib et al. [42]	UK	MER	None	1	1127	N	N	N	N	N
Riado Minguez et al. [43]	Spain	MES	None	1	446	N	Y	Y	Y (150)	N
Sekercioglu et al. [44]	Canada	SSu	None	5	16	N	Y	N	N	N
Shinohara et al. [45]	Japan	MEI, SR	None	1	60	Y	Y	N	N	N
Sims et al. [46]	USA	MES	MES, Su	1	37	Y	N	N	N	N
Storz-Pfennig [47]	Germany	MEA	None	1	13	N	N	N	N	N
Tedesco et al. [48]	Italy	MEEE	None	1	244	N	N	N	N	N
Tsujimoto et al. [49]	Japan	MES	None	3	326	N	N	Y	N	N
Tsujimoto et al. [50]	Japan	MES	None	1	284	N	Y	Y	N	N
Umberham et al. [51]	USA	MER	None	2	265	N	Y	N	N	N
von Niederhausern et al. [52]	Switzerland	SSu	None	2	47	N	Y	N	N	N
Wallach et al. [53]	USA	MESu	MESu	3	64	N	N	N	N	N
Yepes-Nunez et al. [54]	Canada	SSu	None	3	42	N	Y	N	N	N
Yu et al. [55]	Taiwan	LSu	LSu of SRs	1	29	N	Y	N	N	N
Zhang et al. [56]	Canada	SSu	None	4	60	N	Y	N	N	N
Zhang et al. [57]	Canada	SSu	None	1	200	N	Y	Y	Y (200)	Y

LSu literature survey, *M* methods section, *MEA* meta-epidemiological analysis, *MEE* meta-epidemiological evidence, *MEEE* meta-epidemiologic empirical evaluation, *MEI* meta-epidemiological investigation, *MER* meta-epidemiological review, *MES* meta-epidemiological study, *MESc* comparative meta-epidemiological study, *MESu* meta-epidemiological survey, *MO* methodological overview, *MR* methodological review, *N* no, *SR* systematic review, *SSu* systematic survey, *Su* survey, *T/A* title or abstract section, *UK* United Kingdom, *USA* United States of America, *Y* yes

Abbreviations
CI: Confidence interval; CONSORT: Consolidated Standards of Reporting Trials; EQUATOR: Enhancing the QUAlity and Transparency Of health Research; IQR: Interquartile range; MeSH: Medical Subject Headings; MR: Methodological review or review of methods; *N/n*: Sample size; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SD: Standard deviation

Acknowledgements
We thank Professor Lehana Thabane for his editorial support and helpful feedback on drafts of this paper.

Authors' contributions
LM and DOL conceptualized and designed the study. DOL, AL, and LM collected, analyzed, and interpreted the data. DOL wrote the first draft of the manuscript. DOL, LM, and AL contributed to the critical revision of the manuscript, and all authors read and approved the final version.

Funding
No funding sources supported the conduct of this research.

Availability of data and materials
In addition to the supplementary information files which are included in this published article, the dataset generated and analyzed during the current study is available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable—ethics committee approval and consent to participate was not required as this study used previously published data.

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada. [2]Faculty of Health Sciences, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada. [3]Biostatistics Unit, Father Sean O'Sullivan Research Centre, St. Joseph's Healthcare Hamilton, Hamilton, ON L8N 4A6, Canada.

Received: 24 February 2019 Accepted: 20 December 2019

References

1. Guyatt GH, Rennie D, Meade MO, Cook DJ, American Medical Association, JAMA Network. Users' guides to the medical literature: a manual for evidence-based clinical practice. In: Third edition edn, vol. xxv; 2015. p. 697.
2. Hulley SB. Designing clinical research. 4th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2013.
3. 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;1:MR000031.
4. Ghimire S, Kyung E, Kang W, Kim E. Assessment of adherence to the CONSORT statement for quality of reports on randomized controlled trial abstracts from four high-impact general medical journals. Trials. 2012;13:77.
5. Ma J, Akhtar-Danesh N, Dolovich L, Thabane L, investigators C. Imputation strategies for missing binary outcomes in cluster randomized trials. BMC Med Res Methodol. 2011;11:18.
6. Mbuagbaw L, Rochwerg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, Guyatt GH. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev. 2017;6(1):79.
7. Mbuagbaw L, Thabane M, Vanniyasingam T, Borg Debono V, Kosa S, Zhang S, Ye C, Parpia S, Dennis BB, Thabane L. Improvement in the quality of abstracts in major clinical journals since CONSORT extension for abstracts: a systematic review. Contemp Clin Trials. 2014;38(2):245–50.
8. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
9. The EQUATOR Network: Enhancing the QUAlity and Transparency Of health Research [http://www.equator-network.org]
10. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med. 2010;7(2):e1000217.
11. Jin Y, Sanger N, Shams I, Luo C, Shahid H, Li G, Bhatt M, Zielinski L, Bantoto B, Wang M, 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.
12. Moher D, Jones A, Lepage L, Group C. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001;285(15):1992–5.
13. Umbrella reviews: evidence synthesis with overviews of reviews and meta-epidemiologic studies. New York: Springer Berlin Heidelberg; 2016.
14. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. Evidence-based medicine. 2017;22(4):139–42.
15. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
16. Ivers NM, Taljaard M, Dixon S, Bennett C, McRae A, Taleban J, Skea Z, Brehaut JC, Boruch RF, Eccles MP, et al. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. BMJ. 2011;343:d5886.
17. Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. BMJ. 2012;344:e4178.
18. Shanthanna H, Kaushal A, Mbuagbaw L, Couban R, Busse J, Thabane L. A cross-sectional study of the reporting quality of pilot or feasibility trials in high-impact anesthesia journals. Can J Anaesth. 2018;65(11):1180–95.
19. Tanniou J, van der Tweel I, Teerenstra S, Roes KC. Subgroup analyses in confirmatory clinical trials: time to be specific about their purposes. BMC Med Res Methodol. 2016;16:20.
20. Puljak L. Reporting checklist for methodological, that is, research on research studies is urgently needed. J Clin Epidemiol. 2019;112:93.
21. Long L. Routine piloting in systematic reviews--a modified approach? Syst Rev. 2014;3:77.
22. Newman M, Van den Bossche P, Gijbels D, McKendree J, Roberts T, Rolfe I, Smucny J, De Virgilio G. Campbell Collaboration Review Group on the effectiveness of problem-based L: responses to the pilot systematic review of problem-based learning. Med Educ. 2004;38(9):921–3.
23. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276–82.
24. Cebul RD, Hershey JC, Williams SV. Using multiple tests: series and parallel approaches. Clin Lab Med. 1982;2(4):871–90.
25. Abramson JH. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. Epidemiol Perspect Innov. 2011;8(1):1.
26. WordArt.com: Word cloud art creator [https://www.wordart.com]
27. Abbade LPF, Wang M, Sriganesh K, Jin Y, Mbuagbaw L, Thabane L. The framing of research questions using the PICOT format in randomized controlled trials of venous ulcer disease is suboptimal: a systematic survey. Wound Repair Regen. 2017;25(5):892–900.
28. Abdul-Khalek RA, Darzi AJ, Godah MW, Kilzar L, Lakis C, Agarwal A, Abou-Jaoude E, Meerpohl JJ, Wiercioch W, Santesso N, et al. Methods used in adaptation of health-related guidelines: A systematic survey. J Glob Health. 2017;7(2):020412.
29. Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. Am J Phys Med Rehabil. 2017;96(1):34–44.
30. Bolvig J, Juhl CB, Boutron I, Tugwell P, Ghogomu EAT, Pardo JP, Rader T, Wells GA, Mayhew A, Maxwell L, et al. Some Cochrane risk of bias items are not important in osteoarthritis trials: a meta-epidemiological study based on Cochrane reviews. J Clin Epidemiol. 2017.
31. Chase Kruse B, Matt Vassar B. Unbreakable? An analysis of the fragility of randomized trials that support diabetes treatment guidelines. Diabetes Res Clin Pract. 2017;134:91–105.
32. Ebrahim S, Vercammen K, Sivanand A, Guyatt GH, Carrasco-Labra A, Fernandes RM, Crawford MW, Nesrallah G, Johnston BC. Minimally important differences in patient or proxy-reported outcome studies relevant to children: a systematic review. Pediatrics. 2017;139:3.
33. El Dib R, Tikkinen KAO, Akl EA, Gomaa HA, Mustafa RA, Agarwal A, Carpenter CR, Zhang Y, Jorge EC, Almeida R, et al. Systematic survey of randomized trials evaluating the impact of alternative diagnostic strategies on patient-important outcomes. J Clin Epidemiol. 2017;84:61–9.

34. Ge L, Tian JH, Li YN, Pan JX, Li G, Wei D, Xing X, Pan B, Chen YL, Song FJ, et al. Association between prospective registration and overall reporting and methodological quality of systematic reviews: a meta-epidemiological study. J Clin Epidemiol. 2018;93:45–55.
35. Gorne LD, Diaz S. A novel meta-analytical approach to improve systematic review of rates and patterns of microevolution. Ecol Evol. 2017;7(15):5821–32.
36. Khan M, Evaniew N, Gichuru M, Habib A, Ayeni OR, Bedi A, Walsh M, Devereaux PJ, Bhandari M. The fragility of statistically significant findings from randomized trials in sports surgery: a systematic survey. Am J Sports Med. 2017;45(9):2164–70.
37. Kosa SD, Mbuagbaw L, Debono VB, Bhandari M, Dennis BB, Ene G, Leenus A, Shi D, Thabane M, Valvasori S, et al. Agreement in reporting between trial publications and current clinical trial registry in high impact journals: a methodological review. Contemp Clin Trials. 2017;65:144–50.
38. Kovic B, Zoratti MJ, Michalopoulos S, Silvestre C, Thorlund K, Thabane L. Deficiencies in addressing effect modification in network meta-analyses: a meta-epidemiological survey. J Clin Epidemiol. 2017;88:47–56.
39. Kuriyama A, Takahashi N, Nakayama T. Reporting of critical care trial abstracts: a comparison before and after the announcement of CONSORT guideline for abstracts. Trials. 2017;18(1):32.
40. Manja V, AlBashir S, Guyatt G. Criteria for use of composite end points for competing risks-a systematic survey of the literature with recommendations. J Clin Epidemiol. 2017;82:4–11.
41. Papageorgiou SN, Hochli D, Eliades T. Outcomes of comprehensive fixed appliance orthodontic treatment: A systematic review with meta-analysis and methodological overview. Korean J Orthod. 2017;47(6):401–13.
42. Ratib S, Wilkes SR, Grainge MJ, Thomas KS, Tobinska C, Williams HC. Is there an association between study size and reporting of study quality in dermatological clinical trials? A meta-epidemiological review. Br J Dermatol. 2017;176(6):1657–8.
43. Riado Minguez D, Kowalski M, Vallve Odena M, Longin Pontzen D, Jelicic Kadic A, Jeric M, Dosenovic S, Jakus D, Vrdoljak M, Poklepovic Pericic T, et al. Methodological and reporting quality of systematic reviews published in the highest ranking journals in the field of pain. Anesth Analg. 2017;125(4):1348–54.
44. Sekercioglu N, Al-Khalifah R, Ewusie JE, Elias RM, Thabane L, Busse JW, Akhtar-Danesh N, Iorio A, Isayama T, Martinez JP, et al. A critical appraisal of chronic kidney disease mineral and bone disorders clinical practice guidelines using the AGREE II instrument. Int Urol Nephrol. 2017;49(2):273–84.
45. Shinohara K, Suganuma AM, Imai H, Takeshima N, Hayasaka Y, Furukawa TA. Overstatements in abstract conclusions claiming effectiveness of interventions in psychiatry: a meta-epidemiological investigation. PLoS One. 2017;12(9):e0184786.
46. Sims MT, Checketts JX, Wayant C, Vassar M. Requirements for trial registration and adherence to reporting guidelines in critical care journals: a meta-epidemiological study of journals' instructions for authors. Int J Evid Based Healthc. 2017.
47. Storz-Pfennig P. Potentially unnecessary and wasteful clinical trial research detected in cumulative meta-epidemiological and trial sequential analysis. J Clin Epidemiol. 2017;82:61–70.
48. Tedesco D, Farid-Kapadia M, Offringa M, Bhutta ZA, Maldonado Y, Ioannidis JPA, Contopoulos-Ioannidis DG. Comparative evidence on harms in pediatric RCTs from less developed versus more developed countries is limited. J Clin Epidemiol. 2017.
49. Tsujimoto H, Tsujimoto Y, Kataoka Y. Unpublished systematic reviews and financial support: a meta-epidemiological study. BMC Res Notes. 2017;10(1):703.
50. Tsujimoto Y, Tsujimoto H, Kataoka Y, Kimachi M, Shimizu S, Ikenoue T, Fukuma S, Yamamoto Y, Fukuhara S. Majority of systematic reviews published in high-impact journals neglected to register the protocols: a meta-epidemiological study. J Clin Epidemiol. 2017;84:54–60.
51. Umberham B, Hedin R, Detweiler B, Kollmorgen L, Hicks C, Vassar M. Heterogeneity of studies in anesthesiology systematic reviews: a meta-epidemiological review and proposal for evidence mapping. Br J Anaesth. 2017;119(5):874–84.
52. von Niederhausern B, Schandelmaier S, Mi Bonde M, Brunner N, Hemkens LG, Rutquist M, Bhatnagar N, Guyatt GH, Pauli-Magnus C, Briel M. Towards the development of a comprehensive framework: qualitative systematic survey of definitions of clinical research quality. PLoS One. 2017;12(7):e0180635.
53. Wallach JD, Sullivan PG, Trepanowski JF, Sainani KL, Steyerberg EW, Ioannidis JP. Evaluation of evidence of statistical support and corroboration of subgroup claims in randomized clinical trials. JAMA Int Med. 2017;177(4):554–60.
54. Yepes-Nunez JJ, Zhang Y, Xie F, Alonso-Coello P, Selva A, Schunemann H, Guyatt G. Forty-two systematic reviews generated 23 items for assessing the risk of bias in values and preferences' studies. J Clin Epidemiol. 2017;85:21–31.
55. Yu T, Enkh-Amgalan N, Zorigt G. Methods to perform systematic reviews of patient preferences: a literature survey. BMC Med Res Methodol. 2017;17(1):166.
56. Zhang Y, Alyass A, Vanniyasingam T, Sadeghirad B, Florez ID, Pichika SC, Kennedy SA, Abdulkarimova U, Zhang Y, Iljon T, et al. A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants with missing outcome data for continuous outcomes in randomized controlled trials. J Clin Epidemiol. 2017;88:67–80.
57. Zhang Y, Florez ID, Colunga Lozano LE, Aloweni FAB, Kennedy SA, Li A, Craigie S, Zhang S, Agarwal A, Lopes LC, et al. A systematic survey on reporting and methods for handling missing participant data for continuous outcomes in randomized controlled trials. J Clin Epidemiol. 2017;88:57–66.
58. In: Oxford Dictionaries: English (online). Oxford: Oxford University Press; 2018.
59. Page MJ, Altman DG, McKenzie JE, Shamseer L, Ahmadzai N, Wolfe D, Yazdi F, Catala-Lopez F, Tricco AC, Moher D. Flaws in the application and interpretation of statistical analyses in systematic reviews of therapeutic interventions were common: a cross-sectional analysis. J Clin Epidemiol. 2018;95:7–18.
60. Page MJ, Shamseer L, Tricco AC. Registration of systematic reviews in PROSPERO: 30,000 records and counting. Syst Rev. 2018;7(1):32.

Publisher's Note

CHAPTER 4: A protocol for the development of a reporting guideline for methodological studies of health research

4.0 Preface

The following pages include the published protocol covering Chapter 4. Reprinted in accordance with the open access copyright license and permission from the author. This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits the author to reuse the material in any non-commercial way they choose under the terms of the license.

Citation:

Lawson, D. O., Puljak, L., Pieper, D., Schandelmaier, S., Collins, G. S., Brignardello-Petersen, R., . . . Mbuagbaw, L. (2020). Reporting of methodological studies in health research: a protocol for the development of the MethodologIcal STudy reportIng Checklist (MISTIC). BMJ Open, 10(12), e040478. doi:10.1136/bmjopen-2020-040478

Open access Protocol

BMJ Open

Reporting of methodological studies in health research: a protocol for the development of the MethodologIcal STudy reportIng Checklist (MISTIC)

Daeria O Lawson ,[1] Livia Puljak ,[2] Dawid Pieper,[3] Stefan Schandelmaier ,[1,4] Gary S Collins,[5] Romina Brignardello-Petersen,[1] David Moher ,[6,7] Peter Tugwell,[8,9] Vivian A Welch,[7,9] Zainab Samaan ,[1] Brett D Thombs ,[10,11] Anders K Nørskov,[12] Janus C Jakobsen ,[12,13] David B Allison,[14] Evan Mayo-Wilson ,[14] Taryn Young,[15] An-Wen Chan,[16] Matthias Briel,[1,4] Gordon H Guyatt,[1] Lehana Thabane,[17,18,19] Lawrence Mbuagbaw [1,20]

To cite: 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

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-040478).

Received 14 May 2020
Revised 05 November 2020
Accepted 23 November 2020

For numbered affiliations see end of article.

Correspondence to
Daeria O Lawson;
lawsod3@mcmaster.ca

ABSTRACT

Introduction Methodological studies (ie, studies that evaluate the design, conduct, analysis or reporting of other studies in health research) address various facets of health research including, for instance, data collection techniques, differences in approaches to analyses, reporting quality, adherence to guidelines or publication bias. As a result, methodological studies can help to identify knowledge gaps in the methodology of health research and strategies for improvement in research practices. Differences in methodological study names and a lack of reporting guidance contribute to lack of comparability across studies and difficulties in identifying relevant previous methodological studies. This paper outlines the methods we will use to develop an evidence-based tool—the MethodologIcal STudy reportIng Checklist—to harmonise naming conventions and improve the reporting of methodological studies.

Methods and analysis We will search for methodological studies in the Cumulative Index to Nursing and Allied Health Literature, Cochrane Library, Embase, MEDLINE, Web of Science, check reference lists and contact experts in the field. We will extract and summarise data on the study names, design and reporting features of the included methodological studies. Consensus on study terms and recommended reporting items will be achieved via video conference meetings with a panel of experts including researchers who have published methodological studies.

Ethics and dissemination The consensus study has been exempt from ethics review by the Hamilton Integrated Research Ethics Board. The results of the review and the reporting guideline will be disseminated in stakeholder meetings, conferences, peer-reviewed publications, in requests to journal editors (to endorse or make the guideline a requirement for authors), and on the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network and reporting guideline websites.

Registration We have registered the development of the reporting guideline with the EQUATOR Network and publicly posted this project on the Open Science Framework (www.osf.io/9hgbq).

Strengths and limitations of this study

- ► To the best of our knowledge, this is the first study to design an evidence-based tool to support the complete and transparent reporting of methodological studies in health research.
- ► This project will help to highlight the current reporting practices of authors of methodological studies to outline a list of key reporting items.
- ► The stakeholders recruited for the consensus study will represent a diverse group of expert health research methodologists including biostatisticians, clinical researchers, journal editors, healthcare providers and reporting guideline developers.
- ► Our study does not incorporate a blinded consensus process and this may impact the flow of discussions during the conference meetings.

INTRODUCTION

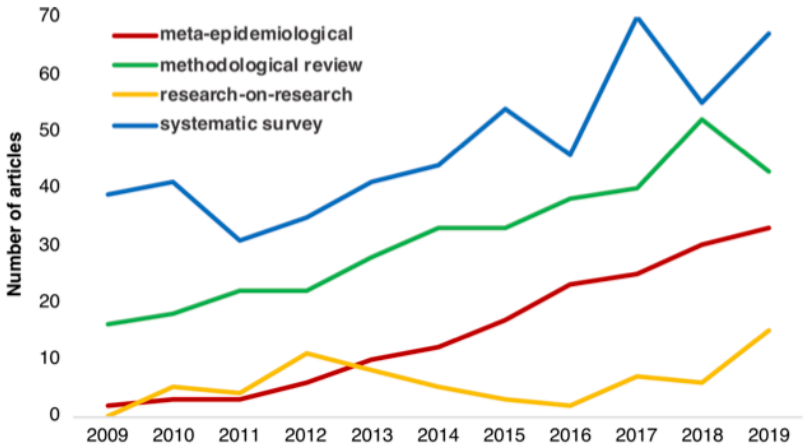
Concerns with the quality and quantity of research have sparked interest in the rapidly evolving field which has been called meta-epidemiology, meta-research or research-on-research.[1–3] This field of research addresses the entire research process, from question development to design, conduct and reporting issues, and most often uses research-related reports (eg, protocols, published manuscripts, registry entries, conference abstracts) as the unit of analysis. These studies may seek to '(1) describe the distribution of research evidence for a specific question; (2) examine heterogeneity and associated risk factors; and (3) control bias across studies and summarise research evidence as appropriate'.[4] For the purpose of this project, we will refer to these research outputs as 'methodological studies', that is,

studies that evaluate the design, conduct, analysis (eg, including bias, statistical plan and methods) or reporting of other studies in health research. This definition does not include statistical methodological studies (eg, studies testing new algorithms or analytical methods, simulation studies) and experimental studies in which the unit of analysis is not a research report. Methodological studies are important because they can identify gaps, biases and inefficiencies in research practices, and propose improvements and solutions.

A PubMed search performed in April 2020 for terms often used to describe methodological studies suggests that the rate of publication of methodological studies has increased over time, illustrated in figure 1.

In the past 20 years, methodological studies have influenced the conduct of health research by informing many popular practices such as double data extraction in systematic reviews[5] ; optimal approaches to conducting subgroup analyses[6] ; and reporting of randomised trials, observational studies, pilot studies and systematic reviews[7–10] to name a few. Methodological studies have played an important role in ensuring that health research is reliable, valid, transparent and replicable. These types of studies may investigate: bias in research,[11 12] quality or completeness of reporting,[13 14] consistency of reporting,[15] methods used,[16] factors associated with reporting practices[17] ; and may provide summaries of other methodological studies[18] and other issues. Methodological studies may also be used to evaluate the uptake of methods over time to investigate whether (and where) practices are improving and allow researchers to make comparisons across different medical areas.[19 20] These studies can also highlight methodological strengths and shortcomings such as sample size calculations in randomised controlled trials,[21 22] quality of clinical prediction models,[23] and spin and over-interpretation of study findings.[24–26] As such, methodological studies promote robust, evidence-based science and help to discard inefficient research practices.[27] A draft conceptual framework of the various categories of methodological studies that we have observed is outlined in figure 2. Broadly, some categories of methodological studies include those investigating: bias and spin, methodological approaches to study design or reporting issues.

Despite the importance of methodological studies, there is no guidance for their reporting. Murad and Wang have suggested a modification to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), a widely used reporting tool that is sometimes used for methodological studies because these studies often use methods that are also used in systematic reviews.[28] Although a modification of PRISMA may work well for the data collection components of some methodological studies, it would fail to appropriately address the many different types of research questions that methodological studies attempt to answer. For example, if researchers were interested in changes in reporting quality of trials since the publication of the CONsolidated Standards Of Reporting Trials guidelines, they could use an interrupted time-series design. Also, methodological studies that include a random sample of research reports,[29] or those structured as before–after designs[19] would be a poor fit for the modified PRISMA tool, which is best suited for studies designed in the style of systematic reviews. Likewise, studies in which the unit of analysis is not the 'study' require more specific guidance (eg, when investigating multiple subgroup analyses or multiple outcomes within the same study).[30] Thus, guidelines for transparent reporting of methodological studies are needed, and this need is widely acknowledged in the scientific community.[31 32]



	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
meta-epidemiological	2	3	3	6	10	12	17	23	25	30	33
methodological review	16	18	22	22	28	33	33	38	40	52	43
research-on-research	0	5	4	11	8	5	3	2	7	6	15
systematic survey	39	41	31	35	41	44	54	46	70	55	67

Figure 1 Trends in methodological studies indexed in PubMed from 2009 to 2019.

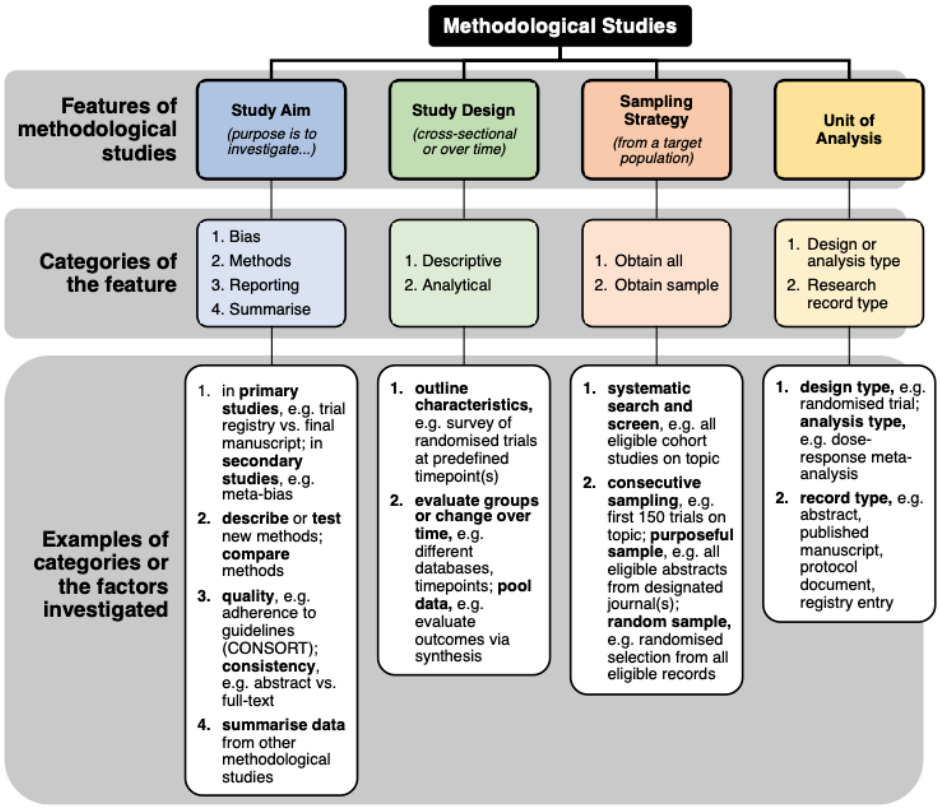


Figure 2 Draft conceptual framework of categories of methodological studies. CONSORT, CONsolidated Standards Of Reporting Trials.

Our work will address two main concerns:

1. There are no globally accepted names for methodological studies, making them difficult to identify. Methodological studies have been called 'methodological review', 'systematic review', 'systematic survey', 'literature review', 'meta-epidemiological study' and many other names. The diversity in names compromises training and educational activities,[33] and it makes it difficult for end-users (eg, clinical researchers, guideline developers) to search for, identify and use these studies.[34 35]
2. The reporting of methodological studies is inconsistent, which may relate to differences in objectives, and to differences in transparency and completeness. That is, some studies may be better reported than others. While the most appropriate approach to reporting will depend on the research question, explicit, user-friendly and consensus-based guidance is needed to ensure that methodological studies are reported transparently and comprehensively.[36]

Aims

The aims of this study protocol are to outline the procedures to define and harmonise the names describing methodological studies, and to develop reporting guidelines for methodological studies in human health research.

METHODS AND ANALYSIS

Study design

We have adopted the strategy for the development of reporting guidelines proposed by Moher *et al*.[37] A visual overview of this approach, highlighting key components of the process, is presented in figure 3. The three parts of the project which will be addressed using the above strategy are outlined in detail below (see online supplemental file for an outline of the data flow informing subsequent parts of the project).

Part 1: methodological review

The objectives of this part are to: (a) identify names used to describe methodological studies, (b) identify the various designs, analysis and reporting features of methodological studies, (c) find any previous reporting guidance and (d) identify methodological study experts.

Search strategy

We developed a search strategy informed by our pilot work[38] targeting health-related sciences and biomedicine databases: Cumulative Index to Nursing and Allied Health Literature, Cochrane Library, Excerpta Medica (Embase), MEDLINE and Web of Science. There will be no limits by publication year, type or language. We will perform searches for authors known to publish in this field, check reference lists of relevant studies, check existing methodological study repositories (Studies Within a Trial

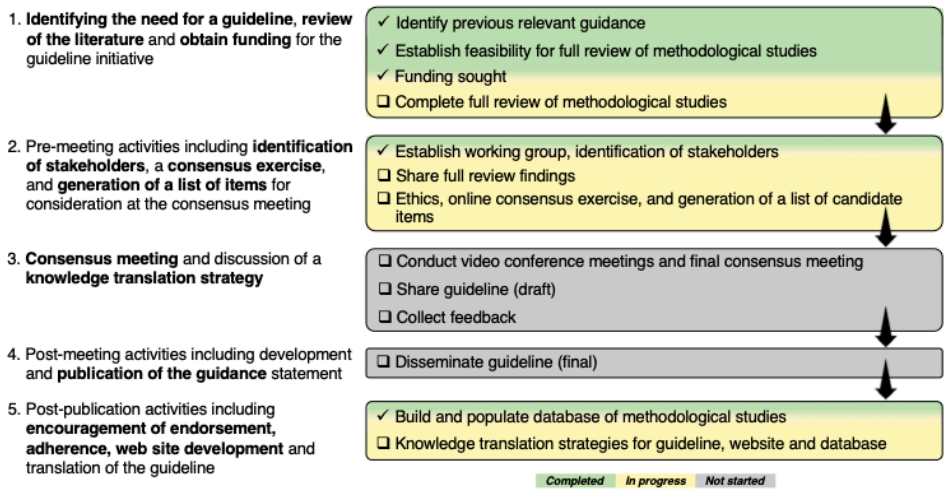


Figure 3 Project overview for the development of reporting guidelines for methodological studies in health research.

and Studies Within a Review), preprints (bioRxiv and medRxiv), set up Google Alerts for keywords (eg, meta-epidemiology, research-on-research) and contact experts (eg, via email, meetings, following relevant journals, subscribing to methods email newsletters including the Methods in Research on Research and the National Institute for Health and Care Excellence groups, and following researchers on social media platforms such as ResearchGate and Twitter) to identify additional methodological studies. We will also check the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) library to identify any published or under development reporting guidance. These approaches are informed by previous work and published literature.[35 38] Two health sciences librarians at the Health Sciences Library (McMaster University) were consulted and reviewed the final search strategy (see online supplemental file) in line with the Peer Review of Electronic Search Strategies framework.[39]

Eligible studies

Studies that investigate methods—design, conduct, analysis or reporting—in other studies of health research in humans will be eligible. The 'other studies' (or research reports) refers to the unit of analysis of the methodological studies (eg, abstracts, cohort studies, randomised trials, registry records, study protocols, systematic reviews). Only published protocols and final reports of studies that investigate methods will be eligible. We will exclude simulation studies, studies testing new statistical methods (ie, there is no specific unit of analysis) and experimental studies of methods (ie, the unit of analysis is not a research report). These sorts of studies either already have reporting guidelines or can be reported in a commentary-style format.

Screening

A team of reviewers led by DOL will screen titles and abstracts independently, in duplicate in Rayyan,[40] and full texts in standardised forms in DistillerSR.[41] Both are online collaborative platforms for screening and reviewing literature. We will measure agreement on screening and study inclusion using Cohen's kappa statistic.[42 43] Any discrepancies between reviewers will be resolved through discussion.

Data extraction

In order to document the current reporting practices, we will extract data from included studies independently, in duplicate based on a standardised data collection form. Key data extraction fields for documenting methodological study features and reporting practices (eg, study design name, databases searched, any guideline use) are outlined in table 1. All data will be compiled in DistillerSR. Any discrepancies between reviewers will be resolved through discussion.

All reviewers will undergo calibration exercises and pilot the screening and data collection forms (25 studies per reviewer). We will incorporate an emergent design in the data collection stage of the review, which is characterised by a flexibility in the methodology, allowing researchers to remain open to modifications.[44] Should any new information that is of interest arise during the full-text screen or data extraction, we will update the data collection form and collect this information for all studies retrospectively and going forward. Any modifications to the present protocol will be reported in the final published review. This iterative approach will allow for the capture of information as new methodological study design features come to light during the full-text screening and data extraction phases. Based on this approach, data extraction will be updated accordingly for previously reviewed studies as needed. For example, we expect to see overlaps in methodological study names, some of which might be attributed to collaborating research groups. There also appear to be similarities in methodological study reporting styles that are borrowed from systematic review[4] or survey study designs, which have both been extensively developed and are omnipresent in health research literature. However, if the current data collection fields, listed in table 1, are

Table 1 Overview of data extraction fields for the review

Section	Data to be collected
Bibliometrics	► Corresponding or last author (first and last name) and contact information (email address). We will first verify whether the corresponding author has academic faculty status, and if not, we will contact the last author ► Country of author ► Publication year ► Study design name in title (verbatim quotation/descriptor) ► Type of article (protocol or final publication, and letter/brief report or full publication) ► Journal
Methods	► Study design name in methods section (verbatim quotation/descriptor) ► Objectives (verbatim quotation) ► Outcomes (verbatim quotation) ► Search strategy reported (yes/no) ► Search time limits and justifications (yes/no and verbatim quotation) ► Databases searched ► Included research report types (eg,randomised trials, systematic reviews, cohort studies) ► Sampling method (where applicable) ► Analysis type (eg, correlation, descriptive, regression, time-series) ► Reporting guidance used and justification (yes/no, name and verbatim quotation) ► Prospective registration and existence of a published protocol (yes/no, where applicable)
Results	► Presence of flow diagrams (yes/no) ► Total records screened and included ► Type of final synthesis performed (qualitative, quantitative, both)
Discussion	► Intended use of findings (verbatim quotation) ► Limitations (verbatim quotation)
Other	► Conflicts of interest (yes/no) ► Funding type (eg, industry, institutional, non-profit) ► Provide access to data (yes/no)

insufficient to capture the nuances of the varieties of methodological studies, we will revise our data collection forms accordingly and collect the data for all studies.

Generation of a list of candidate items

The generation of a list of candidate items will be informed by two sources. First, a list of reporting items will be compiled based on what has been reported by authors of the included studies in the methodological review (eg, flow diagram, search strategy). We will also note the use of any reporting guidance as mentioned by authors (eg, PRISMA, STrengthening the Reporting of OBservational studies in Epidemiology (STROBE)). Each item will be ranked from most frequently reported to those less frequently reported. Second, this list will be presented to expert user stakeholders alongside the proportion of methodological studies that report on each item. Stakeholders will be asked to propose additional relevant items to finalise the list of candidate reporting items for part 2.

Data analysis

We will present the flow of articles retrieved and screened in a study flow diagram, and summarise data in tables with explanatory text. We will provide descriptive statistics, that is, counts (percentage) for categorical data, and means (SD) or medians (IQR) for continuous data. In addition to study names, we will synthesise and tabulate verbatim quotations for the study objectives, outcomes, and intended use of findings to provide context and clarification for methodological study rationales.[45] We will qualitatively group studies into categories based on similarities in reporting features. All statistical analyses will be done in Stata V.15.1.[46] We will identify additional potential stakeholders from the list of authors of included studies.

Part 2: consensus study

This part of the project will consist of consultation with expert user stakeholders in a consensus study. The objectives are to define methodological studies, and outline the recommended study name(s) and best reporting practices. The project steering group (DOL, GHG, LM, LT), which includes members with expertise in health research methods, will oversee the consensus study and development of the reporting guideline.

Identification of stakeholders

The steering group will be responsible for identifying expert user stakeholders based on expertise with methodological studies and expertise with reporting guideline development.[47] Additional stakeholders will be identified from the list of authors (either corresponding or senior, with academic faculty status) of methodological studies from the review. In our selection of stakeholders, we will seek individuals who will be committed to participating and providing feedback for the reporting guideline. We define expert user stakeholders as researchers involved

in the design, conduct, analysis, interpretation or dissemination of methodological studies. Approximately 20–30 stakeholders will be selected (including the protocol authors) as participants in the consensus exercises. We will track response rates to invitations to participate in the consensus study. We will collect participant demographics (eg, country, primary job title, academic rank, and methodological study publication history) to provide insight into the representation in this field of research based on sociocultural factors.

Measuring agreement and achieving consensus

The above definition of methodological studies (ie, studies that evaluate the design, conduct, analysis or reporting of other studies in health research) will be used during the online consensus exercises and video conference meetings. Participants will discuss the following: (a) names for methodological studies, (b) categories of methodological studies and (c) reporting requirements. These three components, outlined in table 2, will be completed electronically through a McMaster Ethics Compliant service, *LimeSurvey* (https://reo.mcmaster.ca/limesurvey) for online surveys.[48]

All video conferences will be facilitated by two investigators (DOL and LM). Stakeholders will be consulted for the development of drafts, elaborations and explanations for specific items. All steering committee members and stakeholders will be required to participate and vote during the consensus meetings. Disagreements will be resolved through discussion, and if no consensus can be reached, the steering committee will convey the recommendations for the stakeholder group to approve. Zoom, or comparable video conferencing software, will be used to allow for the collection of recordings.[49]

Data analysis

Findings from the consensus exercise will be summarised descriptively in tables that include counts (percentage) for categorical data, and means (SD) or medians (IQR) for continuous data. We will measure the levels of agreement (ie, percentage increase in agreement for successive rounds, number of comments made for each successive round and rounds with emergence of new themes) and instability (ie, spread and SD of ranked responses for each item) for each round.[50] After the online exercises, one investigator (DOL) will qualitatively synthesise and code the suggestions for the methodological study names, categories and reporting items into common themes in Dedoose, a qualitative research software.[51] The steering committee will synthesise data from the participant discussions to revise each subsequent draft.

Part 3: reporting guideline

The objectives of this part are to develop, refine, publish and disseminate the reporting guideline for methodological studies. We have registered the development of the reporting guideline— MethodologIcal STudy reportIng Checklist—with the EQUATOR Network.[52] This record may see updates to its name and acronym after deliberations during the consensus study. We will also consider which reporting items are appropriate for different categories of methodological studies. This will include discussions about whether a decision tree may be useful to direct users to other existing reporting guidelines should they be more appropriate for specific categories of methodological studies (eg, STROBE for methodological studies designed as cohort studies). Quantitative and qualitative findings from the consensus study will be incorporated into the final guideline document to include the: (a) recommended methodological study name(s) and categories, (b) recommended checklist with agreed on reporting items, (c) user guide and elaboration (eg, an explanation of why it is important, rationales and an example of how it can be presented in a methodological study), and (d) consensus statement. The draft document will be returned to the steering group and stakeholders to collect additional feedback. The checklist will be tested with end-users for face validity and clarity, and for additional fine-tuning as needed prior to publication. We will distribute the finalised checklist to a group of authors of methodological studies identified from the review (part 1) to assess its usefulness and whether the checklist appropriately captures items relevant to the reporting of methodological studies.[53]

Patient and public involvement

Although patients and the general public are not directly involved in this project, the findings of this research will be relevant to a broad range of knowledge users including methodological study authors, health researchers, methodologists, statisticians and journal editors. We will seek recommendations from investigators for general public members and patients that could be recruited for this project.

ETHICS AND DISSEMINATION

This research has received an exemption (October 2019) from the Hamilton Integrated Research Ethics Board for the consensus study. Ethics committee approval and consent to participate is not required for any other component of this project since only previously published data will be used.

Data deposition and curation

All participant records and data will be stored in MacDrive, a secure cloud storage drive that is privately hosted and based in-house at McMaster University.[54] Only two researchers (DOL and LM) will have direct access to study-related documents and source data. Qualitative data will be promptly coded and transcribed, and all audio files will be encrypted. As part of our knowledge translation (KT) strategy and a consequence of the difficulties we faced in retrieving methodological studies from literature databases during our pilot work, we have developed an open-access database of methodological studies (www.methodsresearch.ca). We will catalogue all included

Table 2 Overview of consensus study activities and expected outputs

Stage	Description of activities to be completed	Expected outputs
Online consensus exercise: categories of methodological studies*	► We will present the **proposed categories of methodological studies** (ie, based on the aim, design, sampling strategy and unit of analysis) with rationale for each. For each category (eg, methodological studies that evaluate study design; methodological studies that evaluate reporting practices) an example of studies that belong in each category will also be presented. An example of potential categories are outlined in figure 2. ► Participants will be asked to rate and comment on the appropriateness of each category on a 3-point ordinal scale: 3—appropriate; 2—somewhat appropriate; 1—inappropriate. ► A validity ratio (VR) will be computed as follows: $VR = \frac{(Ne - N/2)}{(N/2)}$ where ***Ne*** is the number of participants who indicated that the category was appropriate (ie, a rating of '3') and ***N*** is the total number of participants. This ratio will indicate the category that at least half of the participants consider appropriate. The VR will be interpreted based on a table of critical values.[69] For example, for 30 participants (N=30), the critical value is 0.33 (ie, at least 20 participants must deem the category appropriate). Only items based on a critical value greater than the set threshold will be considered further.[70] This approach allows consensus to be achieved remotely and makes decision-making objective.	► List of 'appropriate' categories for methodological studies
Online consensus exercise: name(s) for methodological studies*	► We will present the **names of methodological studies** and for each name (ie, meta-epidemiological study, systematic survey, etc), an example of a study using that name will be provided. ► Participants will be asked to rate the appropriateness of each potential name on a 3-point ordinal scale, and VR will be computed: 3—appropriate; 2—somewhat appropriate; 1—inappropriate	► List of 'appropriate' names for methodological studies
Online consensus exercise: reporting items*	► We will present the **proposed reporting items** and participants will be asked to rate the usefulness of each item on a 3-point ordinal scale, and VR will be computed: 3—essential; 2—maybe essential; 1—not essential ► Participants will be asked to indicate if each reporting item applies to each different methodological study category.	► List of 'essential' reporting items for methodological studies
First video conference meeting (two calls†, 2 hours each)	► Participants will confirm the appropriate name(s) and categories for methodological studies, and agree on reporting items that should be included or excluded, and discuss the rationales behind their selections. All participants will be required to come to a consensus to include an item. ► Meeting minutes and summary of the discussion and decisions will be shared with participants to provide additional feedback after the meeting. ► Based on these discussions and decisions, the steering group will develop a **first draft of the reporting checklist** (eg, with a checkbox to indicate Yes/Reported, No/Not Reported, and a space to indicate on what page the information is reported). ► The checklist will be divided into different reporting sections in a methodological study (eg, Introduction, Methods, Results, Discussion). ► Examples of how to report information for each item will be provided alongside the checklist as part of the **draft user guide**. This will be shared with participants for comment prior to the next meeting.	► First drafts of the: (a) reporting checklist, (b) user guide, (c) recommended methodological study name(s) and categories

Continued

Table 2 Continued

Stage	Description of activities to be completed	Expected outputs
Second video conference meeting (two calls†, 2 hours each)	▲ Participants will agree on a structure and format for the checklist (eg, general layout including appropriate sectioning such as 'Title', 'Abstract' and 'Body' of reports; a decision tree to delineate the category of the methodological study, core items for each category of methodological study, optional items). The group will also review the examples of reporting to be included in the user guide for each reporting item. ▲ Meeting minutes and summary of the discussion and decisions will be shared with participants to provide additional feedback after the meeting. ▲ Based on these discussions and decisions, the steering group will develop a **revised draft of the reporting checklist** and an **elaborated user guide**.	▲ Revised drafts of the: (a) reporting checklist, (b) user guide, (c) recommended methodological study name(s) and categories
Final video conference meeting (4 hours)	▲ Discussion with participants will focus on confirming rationales for the final selected items, and providing examples for each reporting item to be outlined in the consensus statement and elaboration.	▲ Final documents for the: (a) reporting checklist, (b) user guide, (c) recommended methodological study name(s) and categories ▲ Consensus statement and elaboration

*During the online exercises, participants can suggest additional categories, names or items that they wish to discuss during the video conferences.
†Two calls will be scheduled to accommodate stakeholders in Eastern and Western time zones.

studies from the pilot and full reviews on this website such that end-users can easily retrieve these studies. We have also set up a submission portal for researchers to submit their studies to be catalogued in this database. Parallel research by our colleagues will use this database as well as explore the automation of retrieving and indexing methodological studies in a dedicated space.[55] Lastly, we will set up a complementary website to serve as the primary repository for the published reporting guideline document.

Dissemination

We will publish all manuscripts arising from this research and present the findings at conferences. We will set up a complementary website to serve as the primary repository for the published reporting guideline document. The inclusion of knowledge users and representatives from methodology journals and guideline groups on our core study team will aid the wide dissemination of the reporting guideline. We continue to contact journal editors for their endorsement, and encourage researchers to reach out to us about this work, as we have done previously.[34] We will also encourage user feedback to inform future updates of the guideline as needed. These approaches are informed by our collective experience in developing and disseminating health research guidelines.[7 56–60]

DISCUSSION

Our work is contributing to reducing research waste by: (1) making methodological studies transparent through streamlining their reporting; (2) permitting researchers to appraise methodological studies based on adherence to proposed guidelines; (3) allowing end-users of methodological studies to be able to locate inaccessible research in a dedicated database and promoting its continued development; and in doing so (4) allowing end-users of methodological studies to better evaluate and identify issues with study design and reporting that influence patient health, enabling them to apply methodological study evidence to their own research practices. Many methodological studies are done to improve the design, conduct, analysis and reporting of primary and secondary research. We anticipate that, in reviewing this body of evidence on research methods, we will further highlight the importance of studies that aim to improve the design of health research.[61]

Strengths and limitations

We acknowledge that there are inherent challenges in the search and retrieval of studies that lack consistent names, or dedicated indexing in common health research databases. As such, it is plausible that certain methodological studies that use terms not previously identified in the pilot or from our systematic database searches may be missed. To mitigate this limitation, we will (and have already) contact(ed) experts in the field to identify additional studies, and screen references and cite articles

of relevant studies. We have consulted extensively with librarians at the McMaster Health Sciences Library on optimal approaches to capture the maximum number of studies.

The uncertainty in the number of methodological studies that are currently available and published in the literature can present additional logistic and timing constraints to the review component and overall progress of this work. However, given the landscape of methodological studies, we believe it is essential to apply a comprehensive search. To help with the organisation of screening and data extraction, we will use robust systematic review management software (DistillerSR).[41] Further, we have designed all screening and data extraction prompts to ensure consistency and replicability of our work.

Lastly, our study does not incorporate a blinded consensus process and this may impact the flow of discussions during the video conference meetings. We will aim to regulate discussions such that dominant speakers do not steer the discussion and ensure that all participants have a chance to speak. Additionally, we will share summaries of the discussion and decisions after the meetings. This will allow for participants to privately provide any additional written feedback to the steering group that may not have been addressed.

A key strength of this research is the diversity of our study team. We have brought together an international, multidisciplinary team with expertise in consensus activities and guideline development, and research methodology and synthesis. This gives us an advantage in the breadth of feedback and fruitful discussions to be had with a wide array of users of the forthcoming guideline. Given the rise in the conduct of methodological studies, a general call for guidelines in the scientific community, and the number of teams that have reached out to us with interest in participating in this work, we are confident that the guideline will be used. However, we fully acknowledge the factors associated with implementation and use of guidelines, notably journal endorsement of the guidelines, the passage of time and other study level characteristics.[20 62–66] Therefore, our stakeholders include editors from key journals that publish methodological studies such as the *Journal of Clinical Epidemiology*, *BMC Medical Research Methodology*, *BMC Systematic Reviews*, *The Campbell Collaboration*, and *Cochrane*. Stakeholders also include representatives from academic programmes building capacity, at the master's and doctoral level, in conducting methodological research. To encourage better uptake, it has been suggested that researchers should work collaboratively with journals in the prospective design, knowledge translation and evaluation of reporting guidelines,[67] as well as following up on user feedback and incorporating a system to revise the reporting guidelines when necessary.[68] These strategies have been incorporated in our KT plan.

CONCLUSIONS

This research will improve the transparency of reporting of methodological studies, and help streamline their indexing and easier retrieval in literature databases. This work stands to make a substantial impact by informing research reporting standards for studies that investigate the design, conduct, analysis or reporting of other health studies, and thereby improving the transparency, reliability and replicability of health research, and ultimately benefiting patients and decision makers. Future efforts will focus on field-testing the published checklist with authors of methodological studies, gathering feedback from end-users, and optimising and adapting the checklist for different typologies of methodological studies as needed.

Author affiliations
[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
[2]Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia
[3]Institute for Research in Operative Medicine, Witten/Herdecke University, Cologne, Germany
[4]Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University and University Hospital of Basel, Basel, Switzerland
[5]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[6]Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
[7]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[8]School of Epidemiology and Public Health, Faculty of Medicine and Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[9]Bruyère Research Institute, Ottawa, Ontario, Canada
[10]Faculty of Medicine, McGill University, Montreal, Quebec, Canada
[11]Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada
[12]Copenhagen Trial Unit, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[13]Department of Regional Health Research, The Faculty of Heath Sciences, University of Southern Denmark, Odense, Denmark
[14]Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington, Bloomington, Indiana, USA
[15]Centre for Evidence-based Health Care, Division of Epidemiology and Biostatistics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa
[16]Department of Medicine, Women's College Research Institute, University of Toronto, Toronto, Ontario, Canada
[17]Department of Health Research Methods, Evidence, and Impact, and Departments of Paediatrics and Anaesthesia, McMaster University, Hamilton, Ontario, Canada
[18]Biostatistics Unit, Father Sean O'Sullivan Research Centre and Centre for Evaluation of Medicine, Saint Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada
[19]Population Health Research Institute, Hamilton Health Sciences, Hamilton, Ontario, Canada
[20]Biostatistics Unit, Father Sean O'Sullivan Research Centre, Saint Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada

Acknowledgements We would like to thank Denise Smith and Jack Young at the Health Sciences Library at McMaster University for their critical review of the search strategy. We would like to thank Dr Susan Jack at the School of Nursing at McMaster University for sharing her expertise and recommendations on qualitative and mixed-methods research design.

Contributors DOL and LM conceived the idea. DOL, GHG, LM and LT contributed to the design of the study. DOL wrote the first draft of the manuscript. AKN, A-WC, BDT, DBA, DM, DP, EM-W, GHG, GSC, JCJ, LM, LP, LT, MB, PT, RB-P, SS, TY, VAW and

ZS contributed to the refinement of the study methods and critical revision of the manuscript. All authors read and approved the final version of the manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. This project is being carried out as part of a doctoral dissertation by DOL. DOL is supported by an Ontario Graduate Scholarship and the Queen Elizabeth II Graduate Scholarship in Science & Technology, and the Ontario Drug Policy Research Network Student Training Program.

Disclaimer The funders had no influence over the design, data collection, analysis, interpretation, preparation of or decision to publish the manuscript.

Competing interests 'Yes, there are competing interests for one or more authors and I have provided a Competing Interests statement in my manuscript and in the box below'.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Daeria O Lawson http://orcid.org/0000-0002-6487-3367
Livia Puljak http://orcid.org/0000-0002-8467-6061
Stefan Schandelmaier http://orcid.org/0000-0002-8429-0337
David Moher http://orcid.org/0000-0003-2434-4206
Zainab Samaan http://orcid.org/0000-0002-5974-9361
Brett D Thombs http://orcid.org/0000-0002-5644-8432
Janus C Jakobsen http://orcid.org/0000-0002-3642-2120
Evan Mayo-Wilson http://orcid.org/0000-0001-6126-2459
Lawrence Mbuagbaw http://orcid.org/0000-0001-5855-5461

REFERENCES

1 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86–9.
2 Chan A-W, Song F, Vickers A, *et al*. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;383:257–66.
3 Ioannidis JPA, Greenland S, Hlatky MA, *et al*. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383:166–75.
4 Bae J-M. Meta-epidemiology. *Epidemiol Health* 2014;36:e2014019.
5 Buscemi N, Hartling L, Vandermeer B, *et al*. Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol* 2006;59:697–703.
6 Sun X, Briel M, Busse JW, *et al*. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ* 2012;344:e1553.
7 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62:1006–12.
8 Schulz KF, Altman DG, Moher D, *et al*. Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;2010:c332.
9 von Elm E, Altman DG, Egger M, *et al*. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61:344–9.
10 Thabane L, Hopewell S, Lancaster GA, *et al*. Methods and processes for development of a consort extension for reporting pilot randomized controlled trials. *Pilot Feasibility Stud* 2016;2:25.
11 Babic A, Vuka I, Saric F, *et al*. Overall bias methods and their use in sensitivity analysis of Cochrane reviews were not consistent. *J Clin Epidemiol* 2020;119:57–64.
12 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;28:220–32.
13 Croitoru DO, Huang Y, Kurdina A, *et al*. Quality of reporting in systematic reviews published in dermatology journals. *Br J Dermatol* 2020;182:1469–76.
14 Khan MS, Shaikh A, Ochani RK, *et al*. Assessing the quality of Abstracts in randomized controlled trials published in high impact cardiovascular journals. *Circ Cardiovasc Qual Outcomes* 2019;12:e005260.
15 Rosmarakis ES, Soteriades ES, Vergidis PI, *et al*. From conference Abstract to full paper: differences between data presented in conferences and journals. *Faseb J* 2005;19:673–80.
16 Mueller M, D'Addario M, Egger M, *et al*. Methods to systematically review and meta-analyse observational studies: a systematic scoping review of recommendations. *BMC Med Res Methodol* 2018;18:44.
17 Farrokhyar F, Chu R, Whitlock R, *et al*. A systematic review of the quality of publications reporting coronary artery bypass grafting trials. *Can J Surg* 2007;50:266–77.
18 Li G, Abbade LPF, Nwosu I, *et al*. A scoping review of comparisons between Abstracts and full reports in primary biomedical research. *BMC Med Res Methodol* 2017;17:181.
19 Kuriyama A, Takahashi N, Nakayama T. Reporting of critical care trial Abstracts: a comparison before and after the announcement of consort guideline for Abstracts. *Trials* 2017;18:32.
20 Mbuagbaw L, Thabane M, Vanniyasingam T, *et al*. Improvement in the quality of Abstracts in major clinical journals since consort extension for Abstracts: a systematic review. *Contemp Clin Trials* 2014;38:245–50.
21 Armijo-Olivo S, Fuentes J, da Costa BR, *et al*. Blinding in physical therapy trials and its association with treatment effects: a Meta-epidemiological study. *Am J Phys Med Rehabil* 2017;96:34–44.
22 Kovic B, Zoratti MJ, Michalopoulos S, *et al*. Deficiencies in addressing effect modification in network meta-analyses: a meta-epidemiological survey. *J Clin Epidemiol* 2017;88:47–56.
23 Deliu N, Cottone F, Collins GS, *et al*. Evaluating methodological quality of prognostic models including patient-reported health outcomes iN oncologY (EPIPHANY): a systematic review protocol. *BMJ Open* 2018;8:e025054.
24 Kosa SD, Mbuagbaw L, Borg Debono V, *et al*. Agreement in reporting between trial publications and current clinical trial registry in high impact journals: a methodological review. *Contemp Clin Trials* 2018;65:144–50.
25 Schandelmaier S, Chang Y, Devasenapathy N, *et al*. A systematic survey identified 36 criteria for assessing effect modification claims in randomized trials or meta-analyses. *J Clin Epidemiol* 2019;113:159–67.
26 Zhang Y, Flórez ID, Colunga Lozano LE, *et al*. A systematic survey on reporting and methods for handling missing participant data for continuous outcomes in randomized controlled trials. *J Clin Epidemiol* 2017;88:57–66.
27 Ioannidis JPA. Meta-research: why research on research matters. *PLoS Biol* 2018;16:e2005468.
28 Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med* 2017;22:139–42.
29 El Dib R, Tikkinen KAO, Akl EA, *et al*. Systematic survey of randomized trials evaluating the impact of alternative diagnostic strategies on patient-important outcomes. *J Clin Epidemiol* 2017;84:61–9.
30 Tanniou J, van der Tweel I, Teerenstra S, *et al*. Subgroup analyses in confirmatory clinical trials: time to be specific about their purposes. *BMC Med Res Methodol* 2016;16:20.
31 Puljak L. Methodological studies evaluating evidence are not systematic reviews. *J Clin Epidemiol* 2019;110:98–9.
32 Granholm A, Anthon CT, Perner A, *et al*. Transparent and systematic reporting of meta-epidemiological studies. *J Clin Epidemiol* 2019;112:93–5.
33 Rice DB, Moher D. Curtailing the Use of *Preregistration*: A Misused Term. *Perspect Psychol Sci* 2019;14:1105–8.
34 Lawson DO, Thabane L, Mbuagbaw L. A call for consensus guidelines on classification and reporting of methodological studies. *J Clin Epidemiol* 2020;121:110–2.
35 Penning de Vries BBL, van Smeden M, Rosendaal FR, *et al*. Title, Abstract, and keyword searching resulted in poor recovery of articles in systematic reviews of epidemiologic practice. *J Clin Epidemiol* 2020;121:55–61.

36 Puljak L. Reporting checklist for methodological, that is, research on research studies is urgently needed. *J Clin Epidemiol* 2019;112:93.
37 Moher D, Schulz KF, Simera I, *et al*. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217.
38 Lawson DO, Leenus A, Mbuagbaw L. Mapping the nomenclature, methodology, and reporting of studies that review methods: a pilot methodological review. *Pilot Feasibility Stud* 2020;6:1–11.
39 McGowan J, Sampson M, Salzwedel DM, *et al*. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol* 2016;75:40–6.
40 Ouzzani M, Hammady H, Fedorowicz Z, *et al*. Rayyan-a web and mobile APP for systematic reviews. *Syst Rev* 2016;5:210.
41 DistillerSR. in. Ottawa, Canada: evidence partners 2019.
42 Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360–3.
43 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
44 Hesse-Biber SN. *Leavy P: Handbook of emergent methods*. New York: Guilford Press, 2008.
45 Corden A, Sainsbury R. Unit UoYSPR: using Verbatim quotations in reporting qualitative social research: researchers' views: University of York. *Social Policy Research Unit* 2006.
46 Stata Statistical Software. *Release 15.1*. College Station, TX: StataCorp LLC, 2018.
47 Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ* 1995;311:376–80.
48 LimeSurvey. *An open source survey tool*. Hamburg, Germany: Limesurvey GmbH, 2019.
49 Zoom Video Communications. Zoom video communications, Inc. in. San Jose, United States; 2020 2020.
50 Holey EA, Feeley JL, Dixon J, *et al*. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Med Res Methodol* 2007;7:52.
51 Dedoose. *Web application for managing, analyzing, and presenting qualitative and mixed method research data*. Los Angeles, CA: SocioCultural Research Consultants, LLC, 2016.
52 MISTIC - MethodologIcal STudy reportIng Checklist: guidelines for reporting methodological studies in health research. Available: http://www.equator-network.org/library/reporting-guidelines-under-development/reporting-guidelines-under-development-for-other-study-designs/#MISTIC
53 Streiner DL, Norman GR, Cairney J. *Health measurement scales : a practical guide to their development and use*. 15th edn. Oxford: Oxford University Press, 2015.
54 MacDrive. *MacDrive*. Hamilton ON: McMaster University, 2019.
55 SNSF - P3 Research Database: Project 190566. Available: http://p3.snf.ch/Project-190566
56 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
57 Eldridge SM, Chan CL, Campbell MJ, *et al*. Statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud* 2010;2016:64.
58 Guyatt GH, Oxman AD, Vist GE, *et al*. Grade: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
59 Welch V, Petticrew M, Tugwell P, *et al*. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *PLoS Med* 2012;9:e1001333.
60 Wolff RF, Moons KGM, Riley RD, *et al*. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51–8.
61 Hardwicke TE, Serghiou S, Janiaud P, *et al*. Calibrating the scientific ecosystem through Meta-Research. *Annu Rev Stat Appl* 2020;7:11–37.
62 Bastuji-Garin S, Sbidian E, Gaudy-Marqueste C, *et al*. Impact of STROBE statement publication on quality of observational study reporting: interrupted time series versus before-after analysis. *PLoS One* 2013;8:e64733.
63 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.
64 Mills EJ, Wu P, Gagnier J, *et al*. The quality of randomized trial reporting in leading medical journals since the revised consort statement. *Contemp Clin Trials* 2005;26:480–7.
65 Moher D, Jones A, Lepage L, *et al*. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001;285:1992–5.
66 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;6:169–88.
67 Stevens A, Shamseer L, Weinstein E, *et al*. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. *BMJ* 2014;348:g3804.
68 Page MJ, Moher D. Evaluations of the uptake and impact of the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and extensions: a scoping review. *Syst Rev* 2017;6:263.
69 Ayre C, Scally AJ. Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Meas Eval Couns Dev* 2014;47:79–86.
70 Wilson FR, Pan W, Schumsky DA. Recalculation of the Critical Values for Lawshe's Content Validity Ratio. *Meas Eval Couns Dev* 2012;45:197–210.

SUPPLEMENTARY FILE

Timeline of the development of reporting guidelines:

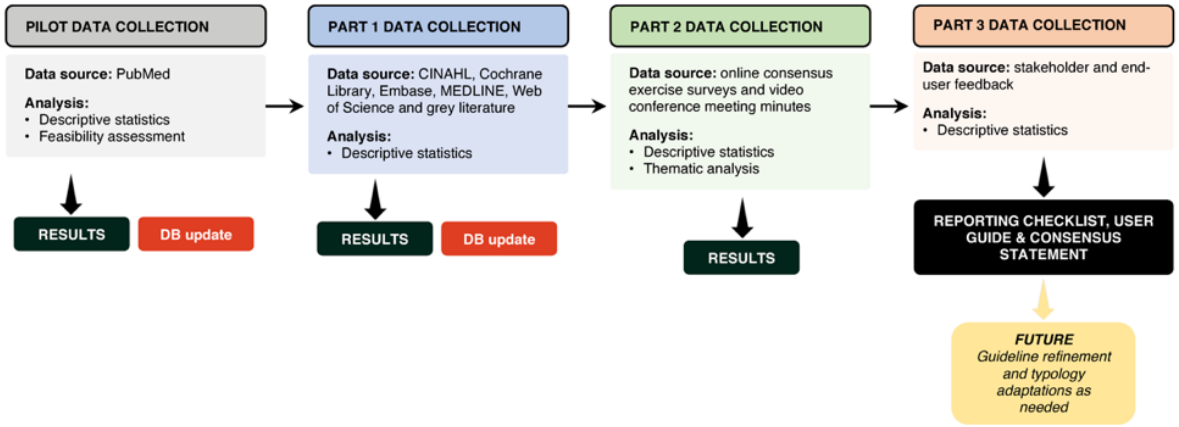
	OBJECTIVES & STEPS	DURATION (MONTHS)	YEAR 1 (2021)				YEAR 2 (2022)			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
1	**Methodological review** of the literature	8	●	●						
2	Development of an **electronic database** for methodological studies	8		●	●					
3	**Consensus study** and development of a **reporting guideline**	6			●	●	●	●		
4	Knowledge translation activities	–	●	●	●	●	●	●	●	●

Q: quarter.

Lawson DO, *et al*. *BMJ Open* 2020; 10:e040478. doi: 10.1136/bmjopen-2020-040478

SUPPLEMENTARY FILE

Flow of data for informing subsequent stages of the project:



CINAHL: Cumulative Index for Nursing and Allied Health Literature, DB: database (www.methodsresearch.ca)

Lawson DO, *et al*. *BMJ Open* 2020; 10:e040478. doi: 10.1136/bmjopen-2020-040478

Ph.D. Thesis – Daeria O. Lawson; McMaster University – Health Research Methodology

SUPPLEMENTARY FILE

Sample search strategy for MEDLINE:

	Concept: names of methodological studies
1	(meta-epidemiolog* OR metaepidemiolog* OR meta-research OR methodolog* analysis OR methodolog* evidence OR methodolog* investigation OR methodolog* literature OR methodolog* overview OR methodolog* report* OR methodolog* review OR methodolog* survey OR methodolog* synthesis OR method* overview OR systematic database review OR systematic literature survey OR systematic survey).mp.
2	(methodolog* study OR method* review OR method* survey)
	Concept: topics in methodological research (i.e. analysis, design and reporting)
3	exp Data Collection/
4	exp Data Interpretation, Statistical/
5	exp Epidemiologic Research Design/
6	exp Nursing Methodology Research/
7	exp Reproducibility of Results/
8	exp Research Design/
9	3 OR 4 OR 5 OR 6 OR 7 OR 8
10	2 AND 9
	Concept: methodological studies that are called 'systematic reviews'
11	systematic review.mp.
12	Cochrane Database of Systematic Reviews.jn
13	11 NOT 12
14	*Data Collection/
15	*Data Interpretation, Statistical/
16	*Epidemiologic Research Design/
17	*Nursing Methodology Research/
18	*Reproducibility of Results/
19	*Research Design/
20	14 OR 15 OR 16 OR 17 OR 18 OR 19
21	13 AND 20
22	1 OR 10 OR 21

CHAPTER 5: A review of methodological studies that evaluate pilot and feasibility studies in health research

The following pages include the unpublished study covering Chapter 5, prepared for submission to an open access journal:

Daeria O. Lawson[1,*], Luis E. Colunga-Lozano[1], Hussein El-Khechen[1], Fang Fang[1], Amy Hutchinson[2], Inam Khan[1], Roger Kou[1], Katie Mellor[3], Agatha Nyambi[1], Nadia Rehman[1], Stefan Schandelmaier[1,4], Yuting Wang[1], Dena Zeraatkar[1], Julian Hirt[5], Maya Stevens-Uninsky[1], Jeslyn Chen[1], Jéssyca Silva[1], Hannah Ewald[6], Behnam Sadeghirad[1], Lehana Thabane[1,7], Lawrence Mbuagbaw[1,7]

1. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON Canada
2. Faculty of Science, McMaster University, Hamilton, ON Canada
3. Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom
4. Clinical Research Empirical Assessment and Recommendations Methods Center, Division of Clinical Epidemiology, Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland
5. Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland
6. University Medical Library, University of Basel, Basel, Switzerland
7. Biostatistics Unit, Father Sean O'Sullivan Research Centre, St. Joseph's Healthcare Hamilton, Hamilton, ON Canada

5.0 Abstract

Background and Objectives:

Methodological research, or research-on-research, aims to evaluate the research process. These studies can be difficult to identify in the literature due to varied names, designs, and no dedicated indexation. This study evaluated methodological studies that investigated pilot and feasibility studies (PAFS) to characterize their nomenclature, reporting practices, and methodology.

Methods:

We searched the Cochrane Database for Systematic Reviews, CINAHL, Embase, MEDLINE, Web of Science, and grey literature sources from inception to June 2023. Studies (final reports or protocols) that investigated the design, conduct, analysis, or reporting of research and included PAFS as the unit of analysis were eligible. Pairs of reviewers extracted data on study design names, methodology, and reporting, summarized descriptively.

Results:

From the initial 19,893 titles and abstracts, 18 articles were included with 16 final reports and two protocols. There were 17 unique study design names used to describe the methodological studies, and five (27.8%) cited the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) as a reporting guide. Two studies published an a priori protocol (12.5%). Thirteen studies evaluated randomized trials and thirteen were specific to a health field (72.2% each). Many studies used descriptive

statistics (16 instances, 45.7%), followed by regression methods (9 instances, 25.7%) to analyze data. The most common aspect of research investigated was reporting practices (13 instances, 54.2%).

Conclusions:

Many methodological studies that assessed PAFS evaluated reporting practices. Diverse study names suggest both heterogeneity in the approach to data synthesis and inconsistencies in naming practices. The reporting style in the methodological studies often mirrored the general layout and flow common to systematic reviews. Although some referred to PRISMA, it is unclear whether this guideline sufficiently addresses all key reporting items for these studies. Future research to streamline the reporting of methodological studies, including on the basis of nomenclature, can help end-users to better identify and access these studies.

Keywords: health research methodology, pilot and feasibility studies, research-on-research, reporting

5.1 Background

The methodological research field increasingly referred to as research-on-research, meta-epidemiology, and meta-research evaluates the practice of research (Bae, 2014; Ioannidis, 2018; Lund et al., 2016; School of Healthcare Enterprise and Innovation, 2023; Sterne et al., 2002). This is a broad and diverse field that is concerned with issues in the entire research process including design, conduct, analysis, and reporting. Over the past three to four decades, these studies have influenced health research by evaluating and exposing deficits in common research practices (Ioannidis, 2018). Examples of deficits include missing or unreliable data (e.g., from erroneous sampling techniques), errors in analyses or interpretation (e.g., use of wrong values, or inappropriate statistical methods), and biases in communicating research or rendering studies non-replicable (Brown, Kaiser, & Allison, 2018; Glasziou et al., 2014). Exposing and correcting errors in research is essential to advancing knowledge about health problems, progressing beyond established ways of thinking (Brown et al., 2018), and reinforcing trust in science (Lawson, Wang, et al., 2022).

Previous research has found that methodological studies can be difficult to identify (Lawson, Leenus, & Mbuagbaw, 2020; Penning de Vries, van Smeden, Rosendaal, & Groenwold, 2020; Puljak, Makaric, Buljan, & Pieper, 2020). This may be partly due to the diversity in names/nomenclature (or lack thereof); diversity in the reporting styles of published manuscripts; and the lack of specific indexing term(s) or a centralized database. It is important to identify methodological studies because they are carried out with the primary goal of improving health research. End-users, including

clinicians, research funders, and health policy makers, need to be able to access methodological studies to apply and put methodological findings into practice.

This study is part of a larger review investigating methodological studies in the health research literature (Lawson, Puljak, et al., 2020). We selected studies that evaluated pilot and feasibility studies (PAFS) as the unit of analysis since there are known methodological discrepancies regarding the design and reporting of PAFS in the literature (Arain, Campbell, Cooper, & Lancaster, 2010; Sim, 2019). PAFS, sometimes referred to as small-scale or proof-of-concept studies, are studies that aim to assess the feasibility of a future, definitive study (Eldridge, Lancaster, et al., 2016; Thabane et al., 2010). The objective of this review was to characterize a subgroup of methodological studies that evaluated PAFS based on their (1) nomenclature, (2) reporting practices including cited reporting guidance, and (3) overall methodology.

5.2 Methods

We published a detailed protocol providing an outline of additional methods for this study (Lawson, Puljak, et al., 2020) and registered our study in the Open Science Framework registry (http://www.osf.io/9hgbq).

Information sources

The search strategy was developed and reviewed with the support of two health sciences librarians, and is in line with Peer Review of Electronic Search Strategies (McGowan et al., 2016). The main concepts were comprised of keywords and MeSH subject headings to target methods-oriented articles based on (1) nomenclature of

methodological studies, (2) topics in methodological research (design, analysis, reporting), and (3) methodological studies that are called "systematic reviews". The latter concept was included to help tease out methods research that is labelled as "systematic reviews" apart from the large body of systematic review literature that is primarily focused on investigating health outcomes. The full search strategy is outlined in Appendix 5.7.1.

The initial search was conducted on August 21, 2019 and it was updated, on June 15, 2020 and on June 7, 2023. Articles were retrieved from the following electronic databases, searched from inception: Cochrane Database for Systematic Reviews (CDSR), Cumulative Index to Nursing and Allied Health Literature (CINAHL) via EBSCOhost, Embase via Ovid, MEDLINE via Ovid, and Web of Science. We searched the following grey literature sources: medRxiv, Studies Within a Trial (SWAT)/Studies Within a Review (SWAR) Repositories, articles suggested by authors and investigators of the MISTIC project, social media accounts of methodologists (ResearchGate and Twitter), Google Alerts, and research newsletters (e.g., EQUATOR Network). There were no timeframe or language limitations to the search.

Eligibility criteria

We included articles (1) reporting on health research on human subjects, and (2) describing research that investigated the design, conduct, analysis, and reporting (including dissemination) of research. Articles that included studies designated as either "pilot" or "feasibility" designs were eligible, either explicitly stated or based on reference lists if PAFS were not explicitly called out by the study authors. Articles were eligible if

the unit of analysis was research records (e.g., PAFS final reports, PAFS abstracts, PAFS protocols). Studies were excluded if they were: (1) methods guidance articles, (2) broad reviews describing methods and overviews of methods in the field (e.g., evidence mapping, opinion, and commentary style articles), (3) described methodological work based on a synthesis of human participant data (e.g., mean ages of included participants) as opposed to methodological aspects of studies, (4) described methodological work to pilot phases of research (e.g., piloting screening in systematic reviews), (5) other studies that were experimental in nature such as SWAT/SWAR, (6) statistical studies (e.g., simulations) and (7) studies reporting PAFS data that was mixed with data from other study designs. Full details for the eligibility criteria checklist are presented in Appendix 5.7.2.

Study selection

We used EndNote X9.3.3, a reference management software (EndNote, 2018), for deduplication of the search results and to retrieve full texts, and Rayyan, a literature review management software (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016) for title and abstract screening. Pairs of reviewers independently screened titles and abstracts for relevance against an a priori eligibility checklist developed for the larger, multi-part review. Raw agreement was calculated for a pilot set of 25 articles between paired reviewers (DOL and AH, HE, HEK, FF, IK, KM, LCL, MSU, NR, SS, YW, and DZ) to ensure validity and reliability of screening. One reviewer (DOL) conducted the final review of conflicts against the same eligibility checklist and if necessary, discussed with a third, senior reviewer (LM).

We used DistillerSR, a literature review management software (DistillerSR, 2023), for full-text screening and data extraction. Study authors and publishers were contacted for full-texts if unavailable via electronic and McMaster Library sources. Reviewers screened full-texts for eligibility in duplicate. Raw agreement was calculated for a pilot set of 25 articles between paired reviewers (DOL and AH, AN, JC, JH, JS, KM, MSU, and RK) to ensure that reviewers were consistent and understood the eligibility criteria before screening the remaining full-texts. Full-texts in non-English languages were reviewed by a single reviewer (trained methodologists and native speakers [JH, JS, LCL, LM, and YW]).

Data extraction

Briefly, the following data was extracted from the included articles: journal, author details, country, study design names (nomenclature), type of article, study objectives and outcomes, search methods, sampling methods, unit of analysis in the study, data analysis methods, reporting guideline use, results presentation, data access, use of study findings, and disclosures. Study design name inconsistencies were logged based on if authors used at least one different name to describe their study in the full text as compared to the name in the title. When authors used clearly abbreviated names (e.g., ‘survey’ in place of ‘systematic survey’, or ‘review’ in place of ‘systematic review’), these were not coded as inconsistencies. Studies were designated as multi-part studies if the authors reported on research comprised of several work packages (e.g., a methodology review and a Delphi survey of participants in one article). Any missing data was indicated as “not reported”, and we did not attempt to contact study authors for

information not available as part of the publication. Full details for all data extraction items, coding schemes, and prompts are presented in Appendix 5.7.2.

We decided that duplicate data extraction was not necessary based on previous experience by our research team (Baldeh et al., 2020), and given the number of eligible studies and detailed guidance manual. Reviewers piloted data extraction on a set of five articles (i.e., as part of the full review which included more eligible articles than the present study), extracted in duplicate by paired reviewers (DOL and AH, AB, JC, JH, JS, KM, MSU, NR, and RK) to ensure that reviewers were consistent and understood all aspects of the data extraction fields before proceeding with the remaining articles. Thereafter, we used a single data extraction with quality control approach. Reviewers extracted data individually and one reviewer (DOL) conducted a quality check by double extracting data for 10% of the articles (i.e., as part of the full review which included more eligible articles than the present study). The same reviewer independently verified all extracted data and amended any discrepancies, or added new data if found.

Data analysis

Data were summarized narratively using explanatory text in tables. Descriptive statistics using counts (percentages) for categorical data, and means (standard deviation, [SD]) and medians (interquartile range, [IQR]) for continuous data were computed. Exact confidence intervals (level = 95%) using the Clopper-Pearson method were constructed. All statistical analyses were conducted in Stata v.15.1 for Macintosh (Stata Statistical Software, 2020).

5.3 Results

Of the 29,717 records identified in the search, after deduplication 19,893 titles and abstracts were screened for relevance. Full-texts of 2,622 potentially eligible articles were assessed, and 18 articles were included in the final analysis with 16 final reports and two protocols. The study flow diagram is outlined in Figure 1. Seventeen articles reported on unique studies and one protocol article described plans for a final report that was also included. One final report described related, follow-up research to a previously published study that was also included. The list of included studies is provided in Appendix 5.7.3.

Description of included studies

The publication of the included articles spanned five years, from 2017 to 2022, with five articles each published in 2019 and 2020. Six were published in *Pilot and Feasibility Studies* (33.3%), followed by two in *BMJ Open* (11.1%). Nine of the articles were published by first authors with a primary affiliation in Canada (50.0%), followed by six based in the United Kingdom (33.3%).

Nomenclature and reporting guideline use

There were 11 varieties of study design names reported in study titles, and 14 in the objectives or methods sections. The names ‘systematic review’ (n = 4), ‘systematic survey’ (n = 3), and ‘review’ (n = 3) appeared most frequently in the title. The names ‘survey’ and ‘systematic survey’ (n = 4, each) appeared most frequently in the objectives or methods section. In total, there were 17 unique study design names attributed to the studies. Inconsistencies in names between sections were present in eight of the articles

(44.4%). In five articles, authors referred to reporting guidelines (27.8%), and all cited the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guideline. In one article, authors reported their protocol according to PRISMA-P but did not mention using PRISMA in their final report of the same study, only that PRISMA-P had been used for the protocol. Data on nomenclature and reporting guideline use is summarized in Table 1.

Methodology

Among the final reports, two reported a published or registered protocol (12.5%). In seven articles, authors reported a full search strategy in the main text (38.9%), and the median (IQR) databases searched was two (1,5). Ten studies included language restrictions (55.6%), and search time limits were applied in 12 (66.7%). Thirteen studies sought to evaluate PAFS randomized controlled trials (72.2%). Various analytical approaches were used including descriptive statistics (n = 16), predictive modelling using regression (n = 9), and narrative synthesis (n = 3). Flow diagrams to illustrate the study title/abstract and full-text screening processes were present in 14 studies (87.5%). Only three studies indicated plans to contact authors for missing data (16.7%). Nine studies were not funded (47.4%), and six were supported by government sources (31.6%). Methodological features of the included articles are summarized in Table 2.

Research intent and use of study findings

Thirteen studies were evaluating research in a specific health field (72.2%). The most common health field was cardiovascular medicine (n = 5), followed by behavioural

science, infectious diseases, nephrology, and obesity-related topics (n = 2, each). The most common study aims were the evaluation of reporting practices (n = 13), and design of studies (n = 8). In 12 articles, authors provided study data access or stated a willingness to do so upon request (75.0%). In seven articles, authors did not explicitly state how the findings would be used (35.0%). The most common stated intents were to inform future research in a broad sense (n = 6), and subsequent methodological research (n = 3). Full details on research intent, data access, and use of study findings are summarized in Table 3. All extracted data for the methodological study objectives and findings is available in Appendix 5.7.4.

5.4 Discussion

The field of methodological research is diverse, and methodological studies evaluating PAFS are no exception, comprising various names, study designs, and reporting styles. In this study, we found that methodological studies evaluating PAFS have been published in the past six years, coinciding with the publication of the Consolidated Standards of Reporting Trials (CONSORT) extension for randomized PAFS in 2016 (Eldridge, Chan, et al., 2016). The publication of CONSORT for PAFS has likely promoted the methodological investigation of PAFS. The extension also led to the inception of the *Pilot and Feasibility Studies* journal, which provides a forum for methodology articles, and a third of the articles in this review (n = 6) were published in this journal. Most often, researchers evaluated reporting practices as the primary

objective. This included investigations such as compliance with reporting guidelines and improvements since the publication of reporting guidelines.

The publication of reporting guidelines has often spurred methodological research investigating study design. Most of the included methodological studies evaluated randomized PAFS (n = 13), and much of the efforts to date on guidance for PAFS has focused on randomized designs. Despite their limitations (e.g., selection bias), nonrandomized PAFS may precede formal randomized trials (Lancaster & Thabane, 2019) and can contribute different types of health information that is complementary to (Booth & Tannock, 2014; Haynes, 2006) or not possible to obtain via randomized trials (e.g., data on disease burden, treatment effectiveness or rare treatment effects, healthcare resource use). In a previous study, we have found that half of the included PAFS studies were based on nonrandomized designs (e.g., cohort, quasi-experimental) including qualitative designs (e.g., focus group) (Lawson, Mellor, et al., 2022). Information from these types of studies is considered to be particularly informative in the field of rehabilitation which is heavily characterized by patient-tailored treatments. In the present study, nonrandomized PAFS were regarded as eligible for inclusion in studies specific to anaesthesiology, nursing, rehabilitation, radiology/medical imaging, and infectious disease research. Future research addressing nonrandomized PAFS can help to improve the methodology in full-scale nonrandomized studies, which is desirable since design-related errors can be avoided in well thought out nonrandomized studies (Goodman, Schneeweiss, & Baiocchi, 2017). In turn, this can help reduce discrepancies and increase agreement between nonrandomized and randomized studies (Boyko, 2013). In the event

formal reporting guidance is developed for nonrandomized PAFS, it can be expected that this will motivate a distinct area of methodological research of nonrandomized PAFS.

Among, the included methodological studies, we found 17 unique study design names. However, the types of methodological features comprising the included articles (e.g., search strategy and database information, process for article selection, flow diagrams to illustrate screening), reflect key items from existing reporting guidelines such as PRISMA. Although no inferential statistics were performed, there was no glaring distinction between the included studies in terms of overall design, except for differences in approaches to study selection and statistical analysis (or qualitative synthesis). Accounting for the diverse end users of this research (e.g., clinicians, methodologists, funding agencies, students), this finding begs the question of whether these studies can be better classified and streamlined, at the very least, by a more unified study design nomenclature. This would also help facilitate access to this body of research.

Among the methodological studies included in this review, PRISMA was the only reporting guideline cited by authors. Although supplemental material of the included articles was not formally evaluated, in a post-hoc scan of the studies that cited PRISMA, none provided the completed PRISMA checklist for their study. This may be due to a combination of (1) appropriateness of PRISMA for methodological studies, (2) lack of journals' endorsement for an alternative, more appropriate guideline although amendments have been proposed (Murad & Wang, 2017), and (3) authors' choice to complete and upload a filled PRISMA checklist with the absence of editorial obligations, especially since the write-up and submission of manuscripts is often already an arduous

process. This finding is not surprising since it is known that journal endorsement is among the key factors for reporting guideline uptake (Jin et al., 2018). This also suggests that a formal reporting guideline development initiative is required to address these issues.

Limitations

This study is not without limitations. It is likely that the search did not capture all relevant methodological studies of PAFS, partly due to the fact that the search strategy was not developed to specifically target pilot and feasibility studies. The different objectives of protocol articles as compared to final reports of research may require separate tabulation and comparison. Since only two protocols were identified in the current study, a meaningful comparison was not possible. In addition, authors of one methodological study evaluating PAFS in nursing conducted a qualitative analysis. It is unclear if this type of methodological study should be grouped together with those that are purely quantitative in nature. Researchers have previously cautioned against the universal application of frameworks from evidence-based medicine such as the procedures in systematic reviews to qualitative studies (Malterud, 2019). Lastly, twelve articles that combined data from non-PAFS studies with data for PAFS were excluded during screening. As a result, these findings should be interpreted with caution and do not represent a full collection of methodological studies evaluating PAFS.

Conclusion

Methodological studies that evaluate pilot and feasibility studies have been

published in the past six years, and most often investigated randomized trials and reporting practices. Authors used many different names to describe their study design, suggesting heterogeneity in the approach to synthesis, but also inconsistency. There are likely key differences between qualitative and quantitative methodological studies that warrants further research. Although PRISMA was cited by some, it was not clear whether this was the most appropriate guideline for reporting such studies. Future research to develop a taxonomy of methodological studies and to assess the appropriateness of existing reporting guidelines can help to streamline the labelling and identification of these studies in the health research literature.

5.5 Tables and Figures

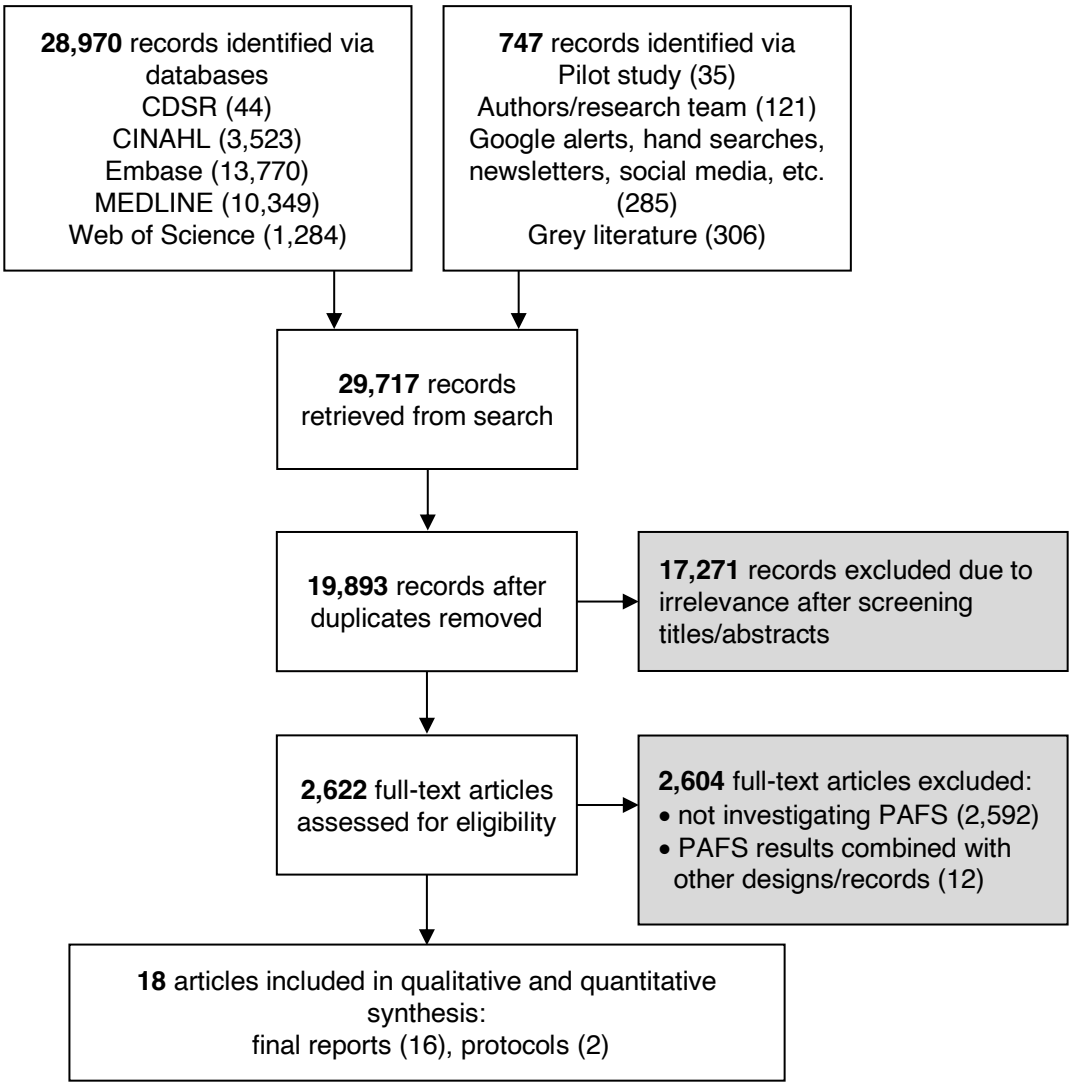


Figure 1. Study flow diagram illustrating screening process and selection of eligible articles

Table 1. Study design nomenclature and reporting guideline use as reported by authors of included studies

Variable	***n* (%)**	**95% CI**
Total, *N*	18 (100)	—
Nomenclature		
Study design name in the title*	21 (100)	—
bibliometric analysis	1 (4.8)	0.6–29.8
descriptive review	1 (4.8)	0.6–29.8
meta-analysis	2 (9.5)	2.2–33.2
meta-epidemiological review	1 (4.8)	0.6–29.8
meta-epidemiological study	2 (9.5)	2.2–33.2
methodological review	1 (4.8)	0.6–29.8
methodological study	1 (4.8)	0.6–29.8
methodological survey	2 (9.5)	2.2–33.2
review	3 (14.3)	4.3–38.0
systematic review	4 (19.1)	6.9–42.9
systematic survey[1]	3 (14.3)	4.3–38.0
Study design name in the objectives/methods section*	27 (100)	—
bibliometric review	1 (3.7)	0.5–23.8
cohort	1 (3.7)	0.5–23.8
descriptive review	1 (3.7)	0.5–23.8
meta-epidemiological investigation	1 (3.7)	0.5–23.8
meta-epidemiological review	1 (3.7)	0.5–23.8

methodologic review	1 (3.7)	0.5–23.8
methodological review	3 (11.1)	3.4–30.6
methodological study	1 (3.7)	0.5–23.8
methodological survey	2 (7.4)	1.7–26.6
narrative synthesis	1 (3.7)	0.5–23.8
review	3 (11.1)	3.4–30.6
survey[2]	4 (14.8)	5.4–34.6
systematic review	3 (11.1)	3.4–30.6
systematic survey[2]	4 (14.8)	5.4–34.6
Are study design names inconsistent between sections – Yes	8 (44.4)	22.7–68.5
Reporting guideline use		
Is there any reference to a specific guideline, checklist, or tool that was used to guide the reporting of the study – Yes	5 (27.8)	11.2–53.9
PRISMA	4 (80.0)	15.2–98.9
PRISMA-P	1 (20.0)	1.1–84.8

CI: confidence interval, PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses, PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols

* studies could contribute more than once

[1] one protocol and final report for the same study account for two instances of "systematic survey"

[2] one protocol and final report for the same study account for two instances each of "survey" and "systematic survey"

Table 2. Methodology of included studies

Variable	***n* (%)**	**95% CI**
Total, *N*	18 (100)	—
Article reporting on a multi-part study – No	18 (100)	—
Refer to or cite a published/registered protocol – Yes[†]	2 (12.5)	2.8–41.7
Search methods:		
Search strategy reported		
Full strategy	7 (38.9)	18.7–63.8
Key terms only	4 (22.2)	8.0–48.6
Key terms and full strategy in supplement	4 (22.2)	8.0–48.6
Supplement only	2 (11.1)	2.5–37.8
Not reported	1 (5.6)	0.7–34.0
Language restrictions applied to the search		
Yes	10 (55.6)	31.5–77.3
No	3 (16.7)	5.0–43.2
Not reported	5 (27.8)	11.2–53.9
Time limits for the search?		
Yes	12 (66.7)	41.1–85.2
No	5 (27.8)	11.2–53.9
Not reported	1 (5.6)	0.7–34.0
Search time limits justified – Yes	7 (58.3)	27.8–83.6
Number of databases/sources searched		
mean (SD)	2.9 (2.0)	1.8–3.9

median (IQR)	2 (1,5)	
Databases/sources searched*		
CENTRAL	3 (6.3)	2.0–18.1
CINAHL	1 (2.1)	0.3–14.0
EBSCOhost	4 (8.3)	3.1–20.6
Embase	5 (10.4)	4.3–23.1
Google Scholar	1 (2.1)	0.3–14.0
MEDLINE	11 (22.9)	13.0–37.2
PsycINFO	2 (4.2)	1.0–15.7
PubMed	12 (25.0)	14.6–39.5
Scopus	3 (6.3)	2.0–18.1
Web of Science	3 (6.3)	2.0–18.1
Others	2 (4.2)	1.0–15.7
Not reported in main text	1 (2.1)	0.3–14.0
Sampling and unit of analysis:		
Types of studies or records that were eligible		
PAFS (nonspecific)	5 (27.8)	11.2–53.9
PAFS RCTs	8 (44.4)	22.7–68.5
PAFS internal/embedded RCTs	1 (5.6)	0.7–34.0
PAFS and associated efficacy/effectiveness trials	1 (5.6)	0.7–34.0
Pilot stepped wedge cluster RCTs, trials/protocols	1 (5.6)	0.7–34.0
PAFS RCTs protocols	2 (11.1)	2.5–37.8
Sample size estimation and basis reported – Yes	2 (11.1)	2.5–37.8

Based on pilot study	1 (50.0)	1.3–98.7
Proportion calculation	1 (50.0)	1.3–98.7
How were studies sampled from the literature?		
All eligible	14 (77.8)	51.4–92.0
All eligible pairs	1 (5.6)	0.7–34.0
Consecutively within journals	1 (5.6)	0.7–34.0
Random	2 (11.1)	2.5–37.8
Analysis:		
Type of analysis or synthesis performed/planned?*		
Critical appraisal of a case study	1 (2.9)	0.4–18.8
Descriptive statistics	16 (45.7)	29.7–62.7
Narrative synthesis	3 (8.6)	0.4–18.8
Parametric statistics	1 (2.9)	2.7–24.2
Pooled – meta-analysis	2 (5.7)	1.4–21.0
Pooled – meta-regression	2 (5.7)	1.4–21.0
Predictive methods – regression	9 (25.7)	13.6–43.2
Qualitative (inductive coding)	1 (2.9)	0.4–18.8
Flow diagram with study or record selection – Yes†	14 (87.5)	58.3–97.2
Researchers planned to contact authors of their included studies for missing data		
Yes	3 (16.7)	5.0–43.2
Yes, for other reasons	1 (5.6)	0.7–34.0
Not reported	14 (77.8)	51.4–92.0
Other features:		

Funding sources that supported conduct of the study*		
Government	6 (31.6)	14.1–56.6
Institutional	1 (5.3)	0.6–32.5
Non-profit organization	1 (5.3)	0.6–32.5
None	9 (47.4)	25.5–70.3
Not reported	2 (10.5)	2.4–36.1
Conflict of interest in relation to study, any author – Yes	2 (11.1)	2.5–37.8

CENTRAL: Cochrane Central Register of Controlled Trials, CI: confidence interval, CINAHL: Cumulative Index to Nursing and Allied Health Literature, IQR: interquartile range, PAFS: pilot and feasibility studies, RCT: randomized controlled trial, SD: standard deviation,

* studies could contribute more than once

† non-protocol articles n=16

Table 3. Research intent and use of findings as reported by authors of included studies

Variable	***n* (%)**	**95% CI**
Total, *N*	18 (100)	—
Assessed research specific to a field of health – Yes*	13 (72.2)	46.1–88.8
Cardiovascular	5 (26.3)	10.7–51.6
Anaesthesiology	1 (5.3)	0.6–32.5
Behavioural	2 (10.5)	2.4–36.1
Infectious diseases	2 (10.5)	2.4–36.1
Nephrology	2 (10.5)	2.4–36.1
Nursing	1 (5.3)	0.6–32.5
Obesity-related	2 (10.5)	2.4–36.1
Physical activity and exercise	1 (5.3)	0.6–32.5
Physiotherapy	1 (5.3)	0.6–32.5
Radiology and medical imaging	1 (5.3)	0.6–32.5
Rehabilitation	1 (5.3)	0.6–32.5
Primary or first listed, methods-related study objective*		
Design of studies	8 (33.3)	17.0–55.0
Conduct of studies	2 (8.3)	1.9–29.5
Analysis of studies	1 (4.2)	0.5–26.5
Reporting of studies	13 (54.2)	33.6–73.4
Do authors provide access to any study data?†		
Yes	12 (75.0)	46.7–91.1
Not reported	4 (25.0)	8.9–53.3

Refer to appendix/supplementary file for additional study results (i.e., not part of main article) – Yes†	9 (56.3)	30.5–79.0
Total studies or records included in analysis/synthesis†		
mean, SD	111.9 (75.9)	71.4–152.3
median, IQR	96 (48,173)	
How do authors plan to use their study findings*		
Education (methods guidance)	1 (5.0)	0.6–31.1
Future research	6 (30.0)	13.4–54.3
Inform a specific study by authors	1 (5.0)	0.6–31.1
Inform editorial/journal activities and policy	1 (5.0)	0.6–31.1
Inform funders/reviewers of studies	1 (5.0)	0.6–31.1
Methods research	3 (15.0)	4.5–39.6
Not reported	7 (35.0)	16.8–59.0

CI: confidence interval, SD: standard deviation,

* studies could contribute more than once

† non-protocol articles n=16

Acknowledgments

We are grateful to the MISTIC project stakeholders and collaborators who have provided feedback on the protocol for this study. We would like to thank Denise Smith and Jack Young at the Health Sciences Library at McMaster University for their support and critical review in the development of the search strategy. We are thankful to Shahrzad Motaghi-Pisheh and Dr. Federico Germini for their help with translating articles for this review.

Ethics

Ethics committee approval and consent to participate was not required for this research because only previously published data were used.

Data access

All original data for this study is available in the published articles of the included studies.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. This project was carried out as part of a doctoral dissertation by DOL. DOL is supported by an Ontario Graduate Scholarship and the Ontario Drug Policy Research Network Student Training Program.

Disclosures

No disclosures or financial conflicts of interest to report.

5.6 References

Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Med Res Methodol, 10*, 67. doi:10.1186/1471-2288-10-67

Bae, J. M. (2014). Meta-epidemiology. *Epidemiol Health, 36*, e2014019. doi:10.4178/epih/e2014019

Baldeh, T., MacDonald, T., Kosa, S. D., Lawson, D. O., Stalteri, R., Olaiya, O. R., . . . Mbuagbaw, L. (2020). More pilot trials could plan to use qualitative data: a meta-epidemiological study. *Pilot Feasibility Stud, 6*(1), 164. doi:10.1186/s40814-020-00712-z

Booth, C. M., & Tannock, I. F. (2014). Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *Br J Cancer, 110*(3), 551-555. doi:10.1038/bjc.2013.725

Boyko, E. J. (2013). Observational research--opportunities and limitations. *J Diabetes Complications, 27*(6), 642-648. doi:10.1016/j.jdiacomp.2013.07.007

Brown, A. W., Kaiser, K. A., & Allison, D. B. (2018). Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proc Natl Acad Sci U S A, 115*(11), 2563-2570. doi:10.1073/pnas.1708279115

DistillerSR. (2023). DistillerSR. Ottawa, Canada: DistillerSR Inc. Retrieved from https://www.distillersr.com/.

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., . . . PAFS Consensus Group. (2016). CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud, 2*, 64. doi:10.1186/s40814-016-0105-8

Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L., & Bond, C. M. (2016). Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. *PLoS One, 11*(3), e0150205. doi:10.1371/journal.pone.0150205

EndNote. (2018). EndNote (Version EndNote X9.3.3). Philadelphia, PA: Clarivate.

Glasziou, P., Altman, D. G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., . . . Wager, E. (2014). Reducing waste from incomplete or unusable reports of biomedical research. *Lancet, 383*(9913), 267-276. doi:10.1016/S0140-6736(13)62228-X

Goodman, S. N., Schneeweiss, S., & Baiocchi, M. (2017). Using Design Thinking to Differentiate Useful From Misleading Evidence in Observational Research. *JAMA, 317*(7), 705-707. doi:10.1001/jama.2016.19970

Haynes, R. B. (2006). *Clinical epidemiology: how to do clinical practice research* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.

Ioannidis, J. P. A. (2018). Meta-research: Why research on research matters. *PLoS Biol, 16*(3), e2005468. doi:10.1371/journal.pbio.2005468

Jin, Y., Sanger, N., Shams, I., Luo, C., Shahid, H., Li, G., . . . Samaan, Z. (2018). Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? - A systematic review of reviews: an update. *J Multidiscip Healthc, 11*, 495-510. doi:10.2147/JMDH.S155103

Lancaster, G. A., & Thabane, L. (2019). Guidelines for reporting non-randomised pilot and feasibility studies. *Pilot Feasibility Stud, 5*, 114. doi:10.1186/s40814-019-0499-1

Lawson, D. O., Leenus, A., & Mbuagbaw, L. (2020). Mapping the nomenclature, methodology, and reporting of studies that review methods: a pilot methodological review. *Pilot Feasibility Stud, 6*, 13. doi:10.1186/s40814-019-0544-0

Lawson, D. O., Mellor, K., Eddy, S., Lee, C., Kim, K. H., Kim, K., . . . Thabane, L. (2022). Pilot and Feasibility Studies in Rehabilitation Research: A Review and Educational Primer for the Physiatrist Researcher. *American journal of physical medicine & rehabilitation, 101*, 372-383. doi:doi:https://dx.doi.org/10.1097/PHM.0000000000001797

Lawson, D. O., Puljak, L., Pieper, D., Schandelmaier, S., Collins, G. S., Brignardello-Petersen, R., . . . Mbuagbaw, L. (2020). Reporting of methodological studies in health research: a protocol for the development of the MethodologIcal STudy

reportIng Checklist (MISTIC). *BMJ Open, 10*(12), e040478. doi:10.1136/bmjopen-2020-040478

Lawson, D. O., Wang, M. K., Kim, K., Eikelboom, R., Rodrigues, M., Trapsa, D., . . . Moher, D. (2022). Lessons from the COVID-19 pandemic and recent developments on the communication of clinical trials, publishing practices, and research integrity: in conversation with Dr. David Moher. *Trials, 23*(1), 671. doi:10.1186/s13063-022-06624-y

Lund, H., Brunnhuber, K., Juhl, C., Robinson, K., Leenaars, M., Dorch, B. F., . . . Chalmers, I. (2016). Towards evidence based research. *BMJ, 355*, i5440. doi:10.1136/bmj.i5440

Malterud, K. (2019). The Impact of Evidence-Based Medicine on Qualitative Metasynthesis: Benefits to be Harvested and Warnings to be Given. *Qual Health Res, 29*(1), 7-17. doi:10.1177/1049732318795864

McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol, 75*, 40-46. doi:10.1016/j.jclinepi.2016.01.021

Murad, M. H., & Wang, Z. (2017). Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med, 22*(4), 139-142. doi:10.1136/ebmed-2017-110713

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Syst Rev, 5*(1), 210. doi:10.1186/s13643-016-0384-4

Penning de Vries, B. B. L., van Smeden, M., Rosendaal, F. R., & Groenwold, R. H. H. (2020). Title, abstract, and keyword searching resulted in poor recovery of articles in systematic reviews of epidemiologic practice. *J Clin Epidemiol, 121*, 55-61. doi:10.1016/j.jclinepi.2020.01.009

Puljak, L., Makaric, Z. L., Buljan, I., & Pieper, D. (2020). What is a meta-epidemiological study? Analysis of published literature indicated heterogeneous study designs and definitions. *J Comp Eff Res, 9*(7), 497-508. doi:10.2217/cer-2019-0201

School of Healthcare Enterprise and Innovation, N. E., Trials and Studies Coordinating Centre. (2023). Research on Research registry and community hub. Retrieved from https://ror-hub.org/

Sim, J. (2019). Should treatment effects be estimated in pilot and feasibility studies? *Pilot Feasibility Stud, 5*, 107. doi:10.1186/s40814-019-0493-7

Stata Statistical Software. (2020). Stata Statistical Software. College Station, TX: StataCorp LLC.

Sterne, J. A., Juni, P., Schulz, K. F., Altman, D. G., Bartlett, C., & Egger, M. (2002). Statistical methods for assessing the influence of study characteristics on

treatment effects in 'meta-epidemiological' research. *Stat Med, 21*(11), 1513-1524. doi:10.1002/sim.1184

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., . . . Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol, 10*, 1. doi:10.1186/1471-2288-10-1

5.7 Appendix

5.7.1 Search strategy

Source	Initial	Update 1	Update 2

I. DATABASES

Cochrane Database of Systematic Reviews				
1	Type > methodology	**38**	**0**	**6**

CINAHL				
Concept: nomenclature of methodological studies				
1	(meta-epidemiolog* OR metaepidemiolog* OR meta-research OR methodolog* analysis OR methodolog* evidence OR methodolog* investigation OR methodolog* literature OR methodolog* overview OR methodolog* report* OR methodolog* review OR methodolog* study OR methodolog* survey OR methodolog* synthesis OR method* overview OR method* review OR method* survey OR systematic database review OR systematic literature survey OR systematic survey)	1,774	140	1,452
Concept: topics in methodological research (i.e., analysis, design and reporting)				
2	(MH "Methodological Research/MT/ST/TD/UT") OR (MH "Research Methodology/MT/ST/TD/UT")	141	1	18
Concept: methodological studies that are called 'systematic reviews'				
3	(MH "Systematic Review/") OR "systematic review"	45,196	3,499	47,228
4	2 AND 3	18	1	4
5	1 OR 2 OR 4	**1,912**	**141**	**1,470**
Limits: exclude MEDLINE records Update 1 from August 2019 to December 2019 (can only specify search by month) Update 2 from January 2020 to June 2023 (can only specify search by month)				

Embase				
Concept: nomenclature of methodological studies				
1	(meta-epidemiolog* OR metaepidemiolog* OR meta-research OR methodolog*	3,720	348	1,759

	analysis OR methodolog* evidence OR methodolog* investigation OR methodolog* literature OR methodolog* overview OR methodolog* report* OR methodolog* review OR methodolog* survey OR methodolog* synthesis OR method* overview OR systematic database review OR systematic literature survey OR systematic survey).mp.			
2	(methodolog* study OR method* review OR method* survey).mp.	10,868	540	2,555
Concept: topics in methodological research (i.e., analysis, design and reporting)				
3	exp Information Processing/	1,507,088	141,795	864,510
4	exp Methodology/	5,583,926	403,089	1,790,966
5	exp Statistical Analysis/	2,216,761	213,678	969,230
6	3 OR 4 OR 5	7,577,177	545,961	2,526,469
7	2 AND 6	4,183	355	1,766
Concept: methodological studies that are called 'systematic reviews'				
8	systematic review.mp.	279,094	43,947	239,473
9	Cochrane Database of Systematic Reviews.jn	13,498	861	2,608
10	8 NOT 9	271,983	43,441	237,230
11	*Information Processing/	34,858	815	2,941
12	*Methodology/	20,226	559	3,078
13	*Statistical Analysis/	18,372	575	1,660
14	11 OR 12 OR 13	71,881	1,919	7,606
15	10 AND 14	1,110	93	418
16	1 OR 7 OR 15	**9,366**	**743**	**3,661**
Update 1 for 2019 (can only specify search by year)				
Update 2 for 2020 to current (can only specify search by year)				

MEDLINE				
Concept: nomenclature of methodological studies				
1	(meta-epidemiolog* OR metaepidemiolog* OR meta-research OR methodolog* analysis OR methodolog* evidence OR methodolog* investigation OR methodolog* literature OR methodolog* overview OR methodolog* report* OR methodolog* review OR methodolog* survey OR methodolog* synthesis OR method* overview OR systematic database review	2,988	302	1,498

	OR systematic literature survey OR systematic survey).mp.			
2	(methodolog* study OR method* review OR method* survey)	7,599	663	2,852
Concept: topics in methodological research (i.e., analysis, design and reporting)				
3	exp Data Collection/	2,054,409	96,056	339,853
4	exp Data Interpretation, Statistical/	55,440	1,114	1,899
5	exp Epidemiologic Research Design/	1,129,276	45,395	159,813
6	exp Nursing Methodology Research/	16,291	70	44
7	exp Reproducibility of Results/	382,553	16,172	63,605
8	exp Research Design/	424,688	15,032	47,509
9	3 OR 4 OR 5 OR 6 OR 7 OR 8	3,139,905	136,904	485,691
10	2 AND 9	2,613	200	976
Concept: methodological studies that are called 'systematic reviews'				
11	systematic review.mp.	154,712	27,793	140,756
12	Cochrane Database of Systematic Reviews.jn	14,335	598	1,686
13	11 NOT 12	140,836	27,266	139,167
14	*Data Collection/	17,926	309	562
15	*Data Interpretation, Statistical/	13,691	419	564
16	*Epidemiologic Research Design/	1,456	30	63
17	*Nursing Methodology Research/	1,428	2	2
18	*Reproducibility of Results/	2,900	88	368
19	*Research Design/	35,321	1,364	6,147
20	14 OR 15 OR 16 OR 17 OR 18 OR 19	69,319	2,116	7,554
21	13 AND 20	1,348	197	657
22	1 OR 10 OR 21	**6,672**	**677**	**3,000**
Update 1 for 2019 (can only specify search by year)				
Update 2 for 2020 to current (can only specify search by year)				

Web of Science (Core Collection, Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH)				
Concept: nomenclature of methodological studies				
1	ALL=("meta-epidemiolog*" OR "metaepidemiolog*" OR "meta-research" OR "methodolog* analysis" OR "methodolog* evidence" OR "methodolog* investigation" OR "methodolog* literature" OR "methodolog* overview" OR "methodolog* report*" OR "methodolog* review" OR "methodolog* study" OR "methodolog* survey" OR "methodolog* synthesis" OR "method* overview" OR "method* review"	12,419	1,277	5,287

	OR "method* survey" OR "systematic database review" OR "systematic literature survey" OR "systematic survey")			
Concept: topics in methodological research (i.e., analysis, design and reporting)				
2	TS=("data analysis" OR "data collection" OR "data interpretation" OR "research design" OR "research method*" OR "research quality" OR "research reporting" OR "statistical analysis" OR "study reporting")	442,732	40,168	175,983
3	1 AND 2	**733**	**105**	**446**
Update 1 for 2019 (can only specify search by year) Update 2 for January 1, 2020 to June 7, 2023				

II. OTHER SOURCES

Research Team				
1	Articles suggested by authors of methodological studies/research team	**156**	n/a	n/a
Google Alerts, Hand Searches and Web Sources				
1	Google Alerts ("meta-epidemiologic" OR "methodological review") and targeted searches (PubMed)	**160**	n/a	n/a
2	Google Alerts ("meta-epidemiologic" OR "methodological review"), experts (e.g., MISTIC project stakeholders), targeted searches (PubMed, ResearchGate, *Journal of Clinical Epidemiology* via Twitter), and newsletters from MiRoR (Methods in Research on Research) and NICE (National Institute for Health and Care Excellence)	n/a	**98**	n/a
Other				
1	Title/abstract screening (i.e., potentially relevant articles as cited in abstracts) and newsletters from EVBRES (Evidence-Based Research Network), EQUATOR (Enhancing the QUAlity and Transparency Of health Research), NICE and RoR (Research on Research registry and hub)	n/a	n/a	**27**
Original spans January 25, 2018 to August 21, 2019 Update 1 spans August 22, 2019 to June 15, 2020 Update 2 spans June 16, 2020 to June 7, 2023				

Grey Literature				
1	medRxiv – Epidemiology/Health Informatics ("meta-epidemiolog*" OR "meta-research" OR "methodolog* review" OR "method* review" OR "systematic survey")	n/a	**3**	**249**
2	SWAT (Studies Within a Trial) and SWAR (Studies Within a Review) Repositories	n/a	**34**	**20**

TOTAL RECORDS	19,037	1,801	8,879
Deduplicated	**13,385**	**840**	**5,668**
Update 1 set of references deduplicated with original set of references Update 2 set of references deduplicated with Update 1 set of references to cover 2020			

5.7.2 Guidance document for screening and data extraction

BACKGROUND

The objective of this review is to define and describe **studies of methods** in the field of health research. These studies will be referred to as **methodological studies** throughout this guidance document. These studies investigate methods in other studies (i.e., meta-research, research-on-research) and for the purpose of this review, we are interested in all methodological studies that investigate methods including the **design, conduct, analysis,** and **reporting** of health studies. In methodological studies, the **research report** (e.g., final publication, abstract, protocol) **is the unit of analysis**.

There is currently 1) no accepted nomenclature, and 2) no guidelines to inform the reporting of methodological studies. As a result, these studies are called many different names. Some of the most common terms used to describe them include variations of *methodological* (e.g., 'methodologic review'), *meta-epidemiological*, *systematic review* or *survey*, *literature review*. It is also common for methodological studies to take on names borrowed from other established fields and study designs (e.g., bibliometric study, literature review, systematic review).

We will conduct a methodological review of methodological studies. The findings from this review will be used to establish a group of stakeholders and develop items for the MethodologIcal STudy reportIng Checklist (MISTIC). The included methodological studies will be catalogued in a public database.

To the best of our knowledge, the landscape of methodological studies has not been previously documented in this extensive way and the scope of the existing literature is not well known. As a result, this review incorporates an emergent design and certain data extraction fields may be updated and/or added as new study features come to light (e.g., based on entries in the 'free-text' fields). This allows us to capture unanticipated data which will also help to develop the conceptual framework for methodological studies.

The key items that will be screened for and extracted from studies include:

- Verify study is a) in the field of health research with humans
- Verify study is b) investigating methods in other studies (i.e., synthesis based on research reports, not human participants)
- Nomenclature used to describe the study
- Methodology used to conduct the study
- If any guidelines have been reported to aid the reporting of the study

Jump to specific guidance:

REVIEW SOFTWARE

I. Rayyan (https://rayyan.qcri.org/welcome) is a web-based systematic review software which will be used for Title and Abstract Screening. An internet connection is required to access the platform.

1. Click on the above link and select 'Get Started' or 'Sign Up' if you do not already have an account. You will receive an automated invitation to the project workbench using the email address that was used to contact you for this study.
2. Upon login, you will see a screen where you can navigate to **Collaboration Reviews** > select **MISTIC REVIEW #** > click 'Show' button. (Note: the '#' will be indicated in your instruction email).
3. You should then see the list of studies in the main window. If not, select 'Search Methods' in the left pane and click on 'Uploaded References'.
4. As you click on each reference there is a button to select **'Include'** / **'Exclude'** / **'Maybe'**.
5. You may enter labels, comments or specify reasons in the text box located beside the above buttons. After entering once, the labels will show up whenever you select 'Reason', and you should not have to enter manually every time but only select from the drop-down.
6. Note: the interface has been setup to highlight some keywords for inclusion (green) and exclusion (red) to aid you in the review process (see Appendix).

II-III. DistillerSR (https://v2dis-prod.evidencepartners.com/) is a web-based systematic review software which will be used for Full Text Screening and Data Extraction. An internet connection is required to access the platform.

1. Click on the above link and enter your assigned username and password that was provided to you in your invitation email.
2. Upon login, you will see the project **MISTIC Review** (or select project from the dropdown in the upper right corner).
3. Select "Unreviewed" from **Level 2** *(Screening)* to begin screening your assigned references. Full texts can be accessed by clicking on the link for the .pdf at the top of the page. This will open the .pdf in a new window/tab (you may need to allow pop-ups in your browser).
4. Select "Unreviewed" from **Level 3** *(Data Extraction)* to begin extracting data from your assigned references. Full texts can be accessed as listed in step 3, above.
5. Note: references may be skipped using the "Skip" button at the top of the screening/data extraction page for each reference.
6. You can refer to the complete data extraction form below (see pages 5-11).

Please contact me (lawsod3@macmaster.ca) if you have any questions about the software interface or experience any difficulties accessing forms, viewing abstracts, or with any of the .pdf files.

SCREENING

LEVEL 1: Title and Abstract

1. Is this study in the field of health research with humans?	· Yes > proceed to #2 · No > Exclude
2. Is this study investigating methods in other studies? Confirm if the study has a methodological focus (i.e., it aims to describe and evaluate or synthesize data about the **design, conduct, analysis** or **reporting** of studies, abstracts of studies, protocols of studies, registrations of studies, etc.). This does NOT include analytical studies comparing different statistical methods (e.g., simulation studies), bibliometric studies (specific to library sciences, often report on metrics and trends in literature), meta-syntheses (a method of synthesizing data specific to qualitative research), or overviews of systematic reviews (a method of synthesizing data in systematic reviews); these would be **ineligible** (see Appendix for a list of definitions). **To avoid missing the capture of any eligible studies**, base your judgment on the study objective and contents of the abstract as opposed to the nomenclature of the study (e.g., 'a systematic review'). For example, some methodological studies may be called 'systematic reviews' by the authors. As another example, 'overviews of systematic reviews' attempt to summarize health data by synthesizing information from multiple systematic reviews on a related health topic. However, a methodological study may be called an 'overview of systematic reviews' although the aim is to summarize methods in systematic reviews; these would be **eligible** for inclusion.	· Yes > Include · No > Exclude · Unable to tell > Maybe

Avoid selecting 'Maybe' as a reason unless it is absolutely necessary (e.g., abstract or title is missing; OR impossible to determine the eligibility based on the above criteria from the abstract or title alone).

Go back to top

LEVEL 2: Full Text

1. Is this study investigating methods in other studies of health research with humans? Confirm if the study has a methodological focus. That is, it aims to describe and evaluate or synthesize data about the **design, conduct, analysis** or **reporting** in research reports such as final study publications, abstracts of studies, protocols of studies, registrations of studies, etc. Both protocol articles and final reports are eligible. If the methodological review is not the primary focus of the study (e.g., a component of a larger Delphi study) but it is described in the article to a reasonable extent it is considered eligible to allow us to evaluate the reporting of that aspect of the research. **Base your judgment on the study objective as opposed to the nomenclature of the study.** It may be helpful to consider whether the included research reports (e.g., randomized trials, systematic reviews, and so on) in the study are being investigated as is. If the aim of the study is to compare methods from a statistical perspective where the unit of analysis is not the research report (e.g., simulation studies, experimental studies such as those comparing techniques by re-analyzing patient data from multiple studies, or testing new algorithms and statistical methods), then this is a different type of methodological study, outside the scope of this review.	· Yes > Include · No > Exclude
2. Is this study investigating pilot or feasibility studies (i.e., unit of analysis)? Prompt is for sub-study only.	· Yes > Include · No > Exclude
3. If no, select a reason for excluding the study	· Reasons 1-4 (5) Other [free text]

Reasons 1-4:

1) Not health research with humans
2) Not investigating methods (e.g., unit of analysis is not a type of research report)
3) Different type of methodological study (e.g., diagnostic, imaging, or laboratory methods; experimental studies of methods such as SWAT/SWAR*, or statistical methods including simulation studies)
4) No full-text available (e.g., conference abstract)

* SWAT/SWAR: studies within a trial/studies within a review

Go back to top

DATA EXTRACTION

LEVEL 3: Included Articles

For the METHODS, RESULTS and DISCUSSION data extraction sections, **avoid the use of free text fields** as much as possible unless it is absolutely necessary.

Bibliometrics	
1. Article title	· [free text]
2. Year of publication	· ####
3. Journal	· [free text]
4. Copy and paste the list of authors	· [free text]
5. Number of authors Confirm the number of authors if a group is listed as an author. The full list can appear as a footnote or at the end of the article.	· ####
6. First author last name	· [free text]
7. Corresponding author name and e-mail address	· [free text – name] · [free text – email]
8. Country Enter the country of the corresponding author's first listed affiliation. If the corresponding author is not the first author, then enter the country of the corresponding author's affiliation.	· [free text]
9. Is this article describing the final report of a study, an update, or the protocol for a planned study?	· Options 1-3 (4) No
Options 1-3: 1) Final report 2) Update to an earlier version 3) Protocol	
Methods	
10. Specify the study design name as listed in the 'Title' of the article Check all applicable names used by the authors to describe the study design.	· Options 1-58 (59) Other [free text] (60) Not reported
Options 1-58: 1) Bibliometric analysis 2) Bibliometric review 3) Bibliometric study 4) Cochrane methodology review 5) Cochrane review 6) Cochrane review of methodology 7) Cohort study 8) Comparative survey 9) Cross-sectional review 31) Methodology/ic/ical review 32) Methodology/ic/ical scoping review 33) Methodology/ic/ical study 34) Methodology/ic/ical survey 35) Methodology/ic/ical systematic review 36) Methods overview 37) Methods review 38) Methods study 39) Observational review	

10) Cross-sectional study 11) Cross-sectional survey 12) Empirical survey 13) Literature methods review 14) Literature review 15) Literature survey 16) Meta-analysis 17) Meta(-)epidemiology/ic/ical assessment 18) Meta(-)epidemiology/ic/ical investigation 19) Meta(-)epidemiology/ic/ical review 20) Meta(-)epidemiology/ic/ical study 21) Meta(-)epidemiology/ic/ical survey 22) Meta(-)research 23) Meta(-)research study 24) Meta-review 25) Meta(-)study 26) Meta(-)synthesis 27) Methodical review 28) Methodical study 29) Methodology/ic/ical assessment 30) Methodology/ic/ical overview	40) Observational study 41) Overview of systematic reviews 42) Research(-)on(-)research 43) Retrospective cohort 44) Review 45) Scoping literature review 46) Scoping review 47) Simulation study 48) Survey 49) Survey of (the) literature 50) Systematic assessment 51) Systematic literature review 52) Systematic literature survey 53) Systematic methods/methodology review 54) Systematic review 55) Systematic review of methods/methodology 56) Systematic scoping review 57) Systematic study 58) Systematic survey
11. Specify the study design name as listed in the 'Methods' section of the article (includes the Objectives section which typically appears just before the Methods) Check all applicable names used by the authors to describe the study design.	See item #10
12. Do the authors define their study design? If yes, copy and paste the definition statement and any cited references (e.g., "a systematic review is a type of synthesis…")	(1) Yes [free text] (2) No
13. Is this article describing a multi-part study? (e.g., findings of a review AND a simulation study, Delphi study, interviews, etc.) If yes, for all remaining data extraction items only refer to the methodological study component of the article. Refer to the eligibility criteria on page 4 of the guidance manual if needed.	(1) Yes (2) No
14. What is the objective of the study? List all applicable.	· Options 1-5 (6) Other [free text]
Options 1-5: 1) Multi-part study which includes a methodological study 2) To investigate design of studies 3) To investigate conduct of studies (how researchers carry out research, including	4) To investigate analytical approaches in studies 5) To investigate reporting practices (including completeness, bias and discrepancies between sources,

registration, team composition, funding influences, etc.)	dissemination in journals or conferences, to patients, etc.)
15. Specify the primary objective of the study Copy and paste text stating the primary objective. If multiple objectives, report only the first listed.	· [free text]
16. Specify the primary methodological outcome of the study Copy and paste text stating the first listed/primary methodological outcome or question. If the primary outcome is not related to methods (this should not be a common occurrence in this review), but a clinical question for example, and the secondary/additional outcomes are focused on methodological aspects, then copy and paste the first listed methods-relevant outcome.	(1) [free text] (2) Not reported, only objectives stated (3) None, stated as exploratory
17. Is there reference to a published or registered protocol for the study? (i.e., if this is not a protocol) Must be explicitly stated. We will not verify or search for protocols due to the volume of studies. If there is a cited reference for a protocol but it turns out to be another study or a previous version of the review (i.e., if the article is an update), or it is some internal document that is not linked (e.g., as a supplementary file), we will consider these as 'No'. **Cochrane Methodology Reviews:** Protocols will be considered as published based on the presence of the section 'History' and the statement 'protocol first published: issue #, year' in the main article OR if there is a cited protocol (and not just a statement suggesting there was a previous version of the review).	· Options 1-4
Options 1-4: 1) Yes, published or registered (citation or link provided) 2) Yes, as appendix/supplementary file 3) Yes, mentioned but no specific citation information or links provided 4) No	
18. Is a search strategy reported? Confirm that the search strategy is available if authors reference one. This may be listed in an appendix or in the methods section.	· Options 1-5 (6) Other [free text]
Options 1-5: 1) Yes, full strategy/key terms for at least one database 2) Yes, but not in main article (reference supplementary file) 3) Partially, key terms only 4) No, only sources/databases mentioned 5) No search information	
19. Were there language restrictions applied to the search? Must be explicitly stated, do not make assumptions from the search strategy if provided.	(1) Yes (2) No (3) Not reported
20. If yes, state the language(s) the search was restricted to	(1) English

If multiple languages, separate each with a comma.	(2) Other [free text]
21. What databases or sources were searched? List all applicable. Note that some sources may not be "databases" (e.g., Cochrane Library) but may be declared as such by authors.	· Options 1-44 (45) Other [free text] (46) Not reported in main text (appendix) (47) Not reported

Options 1-44:

Databases

1) Allied and Complementary Medicine Database (AMED)
2) bioRxiv (biology preprint server)
3) BIOSIS (biological abstracts)
4) Campbell Collaboration
5) Cumulative Index to Nursing and Allied Health Literature (CINAHL)
6) Cochrane Central Register of Controlled Trials (CENTRAL)
7) Cochrane Database of Systematic Reviews (CDSR)
8) Cochrane Library
9) Cochrane Methodology Register (CMR)
10) Database of Abstracts of Reviews of Effects (DARE)
11) EBSCOhost (platform)
12) Embase (Excerpta Medica Database)
13) Education Resources Information Center (ERIC)
14) Google Scholar
15) Health Technology Assessment Database (via Centre for Reviews & Dissemination)
16) Latin American & Caribbean Health Sciences Literature (LILACS)
17) Library and Information Science Abstracts (LISA)
18) Library, Information Science and Technology Abstracts (LISTA)
19) MEDLINE
20) MEDLINE/PubMed
21) medRxiv (health sciences preprint server)
22) NHS Economic Evaluations Database
23) National Institute for Health and Care Excellence (NICE)
24) Ovid (platform)
25) PreMEDLINE
26) ProQuest
27) PsycINFO
28) PubMed
29) Scopus
30) Web of Science

Grey literature, registries and others

31) Agency for Healthcare Research and Quality (AHRQ)
32) Authors of included studies
33) ClinicalTrials.gov
34) Core Outcome Measures in Effectiveness Trials (COMET)
35) Conference proceedings or abstracts
36) Enhacing the QUAlity and Transparency Of health Research (EQUATOR)
37) Experts in the field
38) Direct searches of journals
39) Previous reviews
40) PROSPERO (international prospective register of systematic reviews)
41) Reference lists
42) Science citation index (SCI)
43) Social science citation index (SSCI)
44) World Health Organization International Clinical Trials Registry (WHO ICTRP)

22. Were there time limits provided for the search? Temporal coverage of databases changes over time and as literature is catalogued, there may be greater coverage of earlier years in more recent publications. Therefore, the years referenced for a database may have been the oldest/most comprehensive at	(1) Yes (2) No (searched from inception) (3) Not reported

the time of the search. We will not evaluate if the reported search timeframe was restricted unless this is explicitly stated by the authors. **Update reviews:** the years searched for updates may only include those years which were not captured in the previous version of the study, and this will NOT be considered a time restriction. We will not look at the original/previous version(s) of the study to verify the presence/absence of time limits if it is not explicitly stated.	
23. If yes, state the search timeframe(s) List as reported in the article, adjoined with a dash where applicable (e.g., 2000-2001). If multiple timeframes were searched, list and separate each with a comma (e.g., Jan 2000-Dec 2001, May 2005-Apr 2006).	· ####-####
24. If yes, are the search time limits justified? Do the authors explicitly provide reasoning for their sampling frame/time limits in the search section? (e.g., investigating practices in the past 10 years, or searched databases from 2010-present because a specific tool was introduced in 2010).	(1) Yes (2) No
25. What types of studies or records were eligible for inclusion? List all applicable as specified by the authors.	· Options 1-46 (47) Other [free text]

Options 1-46:

1) Any design (unspecified)
2) Administrative database studies
3) Case-control studies (including nested)
4) Chart reviews
5) Clinical trials
6) Cohort studies
7) Cross-sectional studies
8) Diagnostic studies
9) Economic evaluations
10) Empirical studies of methods/methods reviews (i.e., like this review)
11) Health technology assessments
12) Mendelian randomization studies
13) Meta-analyses (including IPD, NMA)
14) Meta-analyses of observational/non-randomized studies (including IPD, NMA)
15) Meta-analyses of RCTs (including IPD, NMA)
16) Meta-analyses of various study designs (including IPD, NMA)
17) Mixed-method studies
24) Psychometric studies (e.g., developing or validating questionnaires, scales, etc.)
25) Quasi-experimental studies (before/after or pretest-postest, discrete choice, interrupted time series, q-study, patient preference)
26) Qualitative studies
27) Quantitative studies
28) Rapid reviews
29) RCTs
30) RCTs, abstracts
31) RCTs, protocols
32) RCTs, registrations
33) RCTs, cluster design
34) Scoping reviews
35) Scoping reviews, abstracts
36) Scoping reviews, protocols
37) Simulation studies
38) Systematic reviews
39) Systematic reviews, abstracts
40) Systematic reviews, protocols

18) Modelling studies 19) N-of-1 trials 20) Observational/non-randomized studies 21) Observational/non-randomized studies, abstracts 22) Observational/non-randomized studies, protocols 23) Pilot and feasibility studies 41) Systematic reviews, registrations 42) Systematic reviews of observational/non-randomized studies 43) Systematic reviews of RCTs 44) Overviews of systematic reviews 45) Surveys 46) COVID studies	
26. Does the study include only Cochrane Reviews?	(1) Yes (2) No
27. Does the study compare Cochrane Reviews to others?	(1) Yes (2) No
28. Was a sample size estimation reported?	(1) Yes (2) No
29. If yes, specify the estimated sample size Copy and paste the calculated sample size (i.e., the number of studies/records required).	· #### · Other [free text]
30. If yes, what is the sample size estimation based on? List all applicable.	· Options 1-5 (6) Other [free text] (7) Not reported
Options 1-5: 1) Other similar studies 2) Recommendations in literature or a statistical rule of thumb 3) Precision around a confidence interval 4) Proportion of the estimated 'population' of studies 5) Pilot work or previous research by authors	
31. If yes, specify the method used to estimate the sample size Copy and paste text describing the method used to calculate the sample size.	· [free text] (2) Not reported
32. How were studies sampled from the literature? Specify how eligible studies were selected from the sampling frame? (i.e., from all studies retrieved in their search) If not explicitly stated, can refer to study flow diagram to infer 'all eligible'.	(1) All eligible (2) Random sample (3) Other [free text] (4) Not reported
33. Did the researchers attempt to (or plan to) contact the study authors of their included studies for missing data?	(1) Yes (2) Yes, for other research reasons (not for missing data) (3) Other [free text] (4) Not reported
34. What type of analysis or synthesis was performed/planned?	· Options 1-8 (9) Other [free text] (10) Not reported
Options 1-8: 1) Descriptive (e.g., means, proportions) 4) Qualitative [specify]	

2) Parametric / non-parametric between groups (e.g., Chi-Square, T-test, Mann Whitney U) [specify] 3) Effect measures (e.g., odds ratio, risk ratio, mean difference) [specify] 5) Pooled (e.g., meta-analysis) [specify] 6) Regression [specify] 7) Time-series [specify] 8) Narrative synthesis	
35. Is there any reference to a specific guideline / checklist / tool that was used to guide the reporting of the study? Must be explicitly stated as used to guide reporting (e.g., a study flow diagram labelled as a "PRISMA diagram" would be insufficient to suggest the authors followed the PRISMA reporting guideline).	(1) Yes (2) No
36. If yes, list the reporting guideline(s) List all applicable.	· Options 1-4 (5) Other [free text]
Options 1-4: 1) Joanna Briggs Institute guidance for scoping reviews 2) Meta-analyses Of Observational Studies in Epidemiology (MOOSE) 3) Preferred Reporting Items for Overviews of Reviews (PRIOR) 4) Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 5) Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Protocols (PRISMA-P) 6) modified PRISMA for meta-epidemiological studies (Murad et al.) 7) Strengthening the Reporting of Observational studies in Epidemiology (STROBE)	
37. Did the study assess studies or records that were specific to a field of health?	(1) Yes (2) No
38. If yes, what field of health is this study most relevant to? List all applicable.	· Options 1-65 (see list in Appendix)
Results	
39. Is a flow diagram illustrating the study or record selection provided in the body of the article? (if this is not a protocol)	(1) Yes (2) No
40. Total studies or records screened (if this is not a protocol) Enter the number of records that were screened. Must be explicitly stated in the text or in a flow diagram if provided. **Update reviews:** only extract the number of records screened in the current article (e.g., 200 records screened in the update)	· ###### (2) Not reported
41. Total studies or records included in analysis or synthesis (if this is not a protocol) If multiple types of records were included, list all applicable in the order listed in the article (e.g., 40 abstracts, 75 trial registrations). **Update reviews:** only extract the number of records included in the current article (e.g., 2 new records added in the update)	· [free text]

42. Is there reference to an appendix or supplementary file for any additional study results (i.e., not part of the main article)?	(1) Yes (2) No
43. Study findings (if this is not a protocol) Copy and paste text describing the findings for the first listed/primary methodological outcome (refer to item #15 above) along with any summary statistic if provided (e.g., OR 1.03, 95% CI 0.96, 1.57, p = 0.63). If no primary outcome or no summary statement is provided in the 'Results' section, refer to the 'Discussion' section for the first listed summary of findings statement.	· [free text]
Discussion	
44. Use of study findings Refer to text describing how the authors plan to use the study findings, this must be explicitly stated. List all applicable.	· Options 1-7 (8) Other [free text] (9) Not reported
Options 1-7: 1) Inform future research broadly (e.g., provide recommendations from findings) 2) Inform future research by the authors 3) Inform methodological work (e.g., development of a checklist/tool/guideline) 4) Inform funders of research 5) Inform healthcare (e.g., delivery, policy, or other direct impact on patients) 6) Inform editorial/journal activities and policy 7) Educational purpose (e.g., methods guidance)	
Other Sections	
45. Conflicts of interest Specify if any of the authors have declared any conflicts of interest in relation to the study.	(1) Yes (2) None (COI declaration section is apparent) (3) Not reported
46. Funding sources List all applicable funding sources that directly supported the conduct of the study. Do not include any funding specific to the investigators (e.g., research fellowship awarded to X) unless it is explicitly stated to support the study.	· Options 1-6 (7) Other [free text] (8) Not reported
Options 1-6: 1) None (funding declaration section present) 2) Government 3) Industry 4) Institutional (hospital, university) 5) Non-profit (foundation) 6) Professional associations	
47. Do authors provide access to any of the study data? (if this is not a protocol)	(1) Yes (2) Upon request (3) Not reported
48. OPTIONAL – Is there any mention of COVID in the article?	(1) Yes
49. OPTIONAL – Comments	· [free text]

Enter any additional comments that might be helpful for the assessment of this article (e.g., study limitations, authors mention there is no relevant reporting guideline for their study).	
50. OPTIONAL – References cited Copy and paste any relevant reference(s) that were cited in the study if they are: a) methodological studies that may be eligible for inclusion OR b) other reviews of methodological studies (similar to this review)	• [free text]

Go back to top

APPENDIX

I. Keywords to aid with decision making when reviewing the eligibility of studies. This list is only meant to serve as a guide and does not definitively dictate whether a study should be included or excluded. The decision to include or exclude a study should be carefully made by the reviewer based on the outlined eligibility criteria.

Potentially include	Potentially exclude	
bias	academy	mouse
feasibility*	adaptation	murine
literature review	agriculture/agricultural	mutation
meta(-)epidemiologic	animal(s)	mycology/mycological
meta(-)epidemiological	assay	nuclear
meta(-)epidemiology	business	occupational
meta(-)research	cadaver(s)	outbreak
meta(-)study	canine	patients
methodical review	a case	PCR
methodologic	case report	pharmacokinetic
methodological	case series	pig(s)
methodological review	case study	plant(s)
methodological study	causal/causality	porcine
methodology	cells	poultry
pilot*	chromatography	preclinical
quality	chemical	prevalence
reporting	clinical practice guideline(s)	primate(s)
research on research	conference	professional
systematic	construct validity	prognosis
systematic survey	crop(s)	psychometric
	curriculum	q(-)method
	diagnosis	q(-)methodology
	doctor(s)	q(-)methodological
	dog(s)	q(-)study
	ecology/ecological	rabbit
	education	rat
	efficacy	reagent
	equine	realist review
	farm(s)	recruited
	fellow(s)	residency
	fish	resident(s)
	gene(s)	rodent(s)
	gene expression	safety
	genetic	school(s)
	genome	sequencing
	GIS	SNP
	graduate(s)	soil
	GWAS	specialist(s)
	health professional(s)	spectroscopy
	hypothesis	staff

	in situ in vivo in vitro incidence invertebrate isotope livestock marine marker(s) mice mixed(-)method(s)	student(s) surgical surveillance survey survival tomography transgenic translate/translation validate/validation workplace workshop

* prompt is for sub-study only

Go back to top

II. Definitions and purpose of common, established study designs from which authors of methodological studies may borrow naming conventions. For the purposes of this review, it is easiest to distinguish the below studies by their objectives and unit of analysis (as opposed to the name and methodological approach, which can also vary among the below studies). It is common for authors of methodological studies to interchange and refer to their studies by one or more of the below names.

Naming Convention	Definition and Purpose	Scope / Unit of Analysis
Bibliometric study	Summarize metrics and data related to publications including articles and books and trends in literature (e.g., citations, institution and researcher networks, literature growth, text analysis). This nomenclature is specific to the library and information sciences.	articles, books, and other components as related to literature and publications
Literature survey	Summarize scholarly data and current evidence on a topic, primarily narratively. Flexible in scope (broad or narrow) depending on the researcher's aims. Not specific to health sciences.	various scholarly records and articles
Meta-synthesis	Summarize qualitative health data where aim is to generate a theory about a topic by way of re-interpretation of multiple studies. This nomenclature is specific to qualitative and mixed-methods research. **Relevant reporting guideline:** ENTREQ	primary qualitative studies
Overview of reviews OR **Umbrella review**	Summarize health data by synthesizing information from multiple systematic reviews on a specific health topic. Common if there have been many systematic reviews conducted on a topic with aim of synthesizing the highest level of evidence. **Relevant reporting guideline:** None / PRIOR (in development)	systematic reviews
Scoping review	Identify nature (scope) and extent of available research evidence. Provide an overview on a topic, often broadly, and aim to identify key features or knowledge gaps.	various scholarly records and articles

	Often summarize data narratively, classifying and grouping into themes, or illustratively using figures/charts. May serve as a preliminary assessment of the literature before undertaking a systematic review **Relevant reporting guideline:** PRISMA-ScR	
Systematic review OR **Rapid review** (expedited version of above)	Summarize health data from primary studies where the outcomes address a clinical question. May incorporate a meta-analysis for quantitative data. **Relevant reporting guideline:** PRISMA and extensions	primary quantitative and/or qualitative studies (e.g., randomized trials, cohort studies)
Systematic survey	Identify trends in a population by summarizing data from a physical survey of a population or area (e.g., collection of samples, interviews, surveys). Not specific to health sciences.	population or area of interest (e.g., humans, animals)
Other related designs:		
Studies Within A Trial, Studies Within A Review	Evaluate methodological interventions within ongoing trials or systematic reviews.	human participants (e.g., trialists, patients, systematic reviewers)

Sources (accessed August 9, 2019):
1. Iowa State University Library. https://instr.iastate.libguides.com/c.php?g=49332&p=318077
2. Temple University Libraries. https://guides.temple.edu/systematicreviews
3. University of Toronto Libraries. https://guides.library.utoronto.ca/c.php?g=588615&p=4310109

Go back to top

III. List of health fields for data extraction item #38 (compiled from *the BMJ,* Cochrane Library, and Web of Science health field indices, accessed August 9, 2019).

Health Fields	
1) Allergy & intolerance	35) Neonatal care
2) Anaesthesiology	36) Nephrology
3) Audiology & speech-language pathology	37) Neurology
4) Behavioural sciences	38) Nursing
5) Biostatistics	39) Nutrition, dietetics & metabolism
6) Blood disorders	40) Occupational health
7) Cardiovascular medicine & circulation	41) Orthopaedics
8) Cancer/oncology	42) Paediatrics
9) Clinical practice & health systems	43) Pain & anaesthesia
10) Complementary & alternative medicine	44) Palliative care
11) Critical & intensive care	45) Pathology
12) Dentistry & oral medicine	46) Pharmacology & pharmaceutical medicine
13) Dermatology	
14) Developmental & learning problems	47) Pregnancy & childbirth
15) Disaster medicine & preparedness	48) Prescribing practices
16) Ear, nose & throat/otolaryngology	49) Public health
17) Emergency medicine	50) Psychiatry
18) Endocrinology & metabolism	51) Psychology
19) Epidemiology	52) Radiology & medical imaging
20) Ethics	53) Rehabilitation medicine
21) Eyes & vision/ophthalmology	54) Reproductive health and fertility
22) Family medicine	55) Respiratory medicine
23) Gastroenterology & hepatology	56) Rheumatology
24) Genetics	57) Sexual health
25) Geriatric medicine & gerontology	58) Social work
26) Global health	59) Substance abuse (tobacco, drugs, alcohol)
27) Gynaecology & obstetrics	
28) Haematology	60) Surgery
29) Health care services	61) Toxicology
30) Immunology	62) Transplantation
31) Infectious diseases	63) Trauma & wounds
32) Internal medicine	64) Urology
33) Laboratory medicine	65) Other: [free text]
34) Mental health & psychosocial problems	

Go back to top

5.7.3 List of included studies

1. Baldeh, T., MacDonald, T., Kosa, S. D., Lawson, D. O., Stalteri, R., Olaiya, O. R., . . . Mbuagbaw, L. (2020). More pilot trials could plan to use qualitative data: a meta-epidemiological study. Pilot & Feasibility Studies, 6(1), 164. doi:https://dx.doi.org/10.1186/s40814-020-00712-z
2. Beets, M. W., von Klinggraeff, L., Burkart, S., Jones, A., Ioannidis, J. P. A., Weaver, R. G., . . . Li, X. (2022). Impact of risk of generalizability biases in adult obesity interventions: A meta-epidemiological review and meta-analysis. Obesity Reviews, 23. doi:https://dx.doi.org/10.1111/obr.13369
3. Beets, M. W., Weaver, R. G., Ioannidis, J. P. A., Geraci, M., Brazendale, K., Decker, L., . . . Milat, A. J. (2020). Identification and evaluation of risk of generalizability biases in pilot versus efficacy/effectiveness trials: a systematic review and meta-analysis. International Journal of Behavioral Nutrition & Physical Activity, 17(1), 19. doi:https://dx.doi.org/10.1186/s12966-020-0918-y
4. Bhatt, M., Zielinski, L., Sanger, N., Shams, I., Luo, C., Bantoto, B., . . . Samaan, Z. (2018). Evaluating Completeness of Reporting in Behavioral Interventions Pilot Trials. Research on Social Work Practice, 28(5), 577-584.
5. Charlesworth, M., Klein, A. A., & White, S. M. (2020). A bibliometric analysis of the conversion and reporting of pilot studies published in six anaesthesia journals. Anaesthesia, 75(2), 247-253. doi:10.1111/anae.14817
6. El-Khechen, H. A., Khan, M. I. U., Leenus, S., Olaiya, O., Durrani, Z., Masood, Z., . . . Mbuagbaw, L. (2021). Design, analysis, and reporting of pilot studies in HIV: a systematic review and methodological study. Pilot and Feasibility Studies, 7. doi:https://dx.doi.org/10.1186/s40814-021-00934-9
7. Horne, E., Lancaster, G. A., Matson, R., Cooper, A., Ness, A., & Leary, S. (2018). Pilot trials in physical activity journals: a review of reporting and editorial policy. Pilot Feasibility Stud, 4, 125. doi:10.1186/s40814-018-0317-1
8. Isiguzo, G., Zunza, M., Chirehwa, M., Mayosi, B. M., & Thabane, L. (2017). Quality of abstracts of pilot trials in heart failure: A protocol for a systematic survey. Contemp Clin Trials Commun, 8, 258-263. doi:10.1016/j.conctc.2017.11.004
9. Isiguzo, G. C., Zunza, M., Chirehwa, M., Mayosi, B. M., & Thabane, L. (2018). Quality of pilot trial abstracts in heart failure is suboptimal: a systematic survey. Pilot Feasibility Stud, 4, 107.
10. Kenis, S. F., Abeyakoon, O., Plumb, A. A. O., & Halligan, S. (2020). Do radiological research articles apply the term "pilot study" correctly? Systematic review. Clinical Radiology, 75, 395.e391-395.e395. doi:https://dx.doi.org/10.1016/j.crad.2019.11.010
11. Kosa, S. D., Monize, J., Leenus, A., Leenus, S., Samra, S., Szwiega, S., . . . Thabane, L. (2019). Reporting quality of pilot clinical trials in chronic kidney disease patients on hemodialysis: a methodological survey. Pilot Feasibility Stud, 5, 53. doi:10.1186/s40814-019-0436-3
12. Kristunas, C. A., Hemming, K., Eborall, H., Eldridge, S., & Gray, L. J. (2019). The current use of feasibility studies in the assessment of feasibility for stepped-wedge

cluster randomised trials: a systematic review. BMC Med Res Methodol, 19(1), 12. doi:10.1186/s12874-019-0658-3
13. Lawson, D. O., Mellor, K., Eddy, S., Lee, C., Kim, K. H., Kim, K., . . . Thabane, L. (2022). Pilot and Feasibility Studies in Rehabilitation Research: A Review and Educational Primer for the Physiatrist Researcher. American journal of physical medicine & rehabilitation, 101, 372-383. doi:https://dx.doi.org/10.1097/PHM.0000000000001797
14. Mailhot, T., Goulet, M. H., Maheu-Cadotte, M. A., Fontaine, G., Lequin, P., & Lavoie, P. (2020). Methodological reporting in feasibility studies: a descriptive review of the nursing intervention research literature. Journal of Research in Nursing, 25(5), 460-472. doi:https://dx.doi.org/10.1177/1744987119883404
15. Mbuagbaw, L., Kosa, S. D., Lawson, D. O., Stalteri, R., Olaiya, O. R., Alotaibi, A., & Thabane, L. (2019). The reporting of progression criteria in protocols of pilot trials designed to assess the feasibility of main trials is insufficient: A meta-epidemiological study. Pilot and Feasibility Studies, 5(1).
16. Mellor, K., Eddy, S., Peckham, N., Bond, C. M., Campbell, M. J., Lancaster, G. A., . . . Hopewell, S. (2021). Progression from external pilot to definitive randomised controlled trial: a methodological review of progression criteria reporting. BMJ Open, 11(6), e048178. doi:https://dx.doi.org/10.1136/bmjopen-2020-048178
17. Rosala-Hallas, A., Gamble, C., Blazeby, J., & Williamson, P. R. (2019). A review of current practice in the design and assessment of internal pilots in UK NIHR clinical trials. Trials, 20(1), 571. doi:10.1186/s13063-019-3669-9
18. Scola, L. F. C., Moseley, A. M., Thabane, L., Almeida, M., & Costa, L. (2019). A methodological survey on reporting of pilot and feasibility trials for physiotherapy interventions: a study protocol. BMJ Open, 9(5), e020580. doi:10.1136/bmjopen-2017-020580

Clarifications

a) Beets et al., 2020 and Beets et al., 2022 – articles describe related research with the latter being a follow-up to the earlier study
b) Isiguzo et al., 2017 and Isiguzo et al., 2018 – articles describe the same study where the first publication is a protocol of the final report

5.7.4 Methodological objectives, outcomes, and findings as reported by authors of included studies

Study	Title	Objective	Primary or first listed methods objective	Primary or first listed methods outcome	Findings related to first listed methods objective or outcome (or first listed finding if objective and/or outcome is unspecific)
Baldeh 2020	More pilot trials could plan to use qualitative data: a meta-epidemiological study	design	Therefore, the purpose of this study was to determine how often pilot trials are designed to collect qualitative data and the study characteristics associated with planning to collect qualitative data.	First, counts of the plans to collect qualitative data, as well as the study characteristics, were summarized descriptively in cross tabulations.	Of the 227 included pilot trial protocols, only 92 (40.5%; 95% CI 34.1–47.2) reported plans to collect qualitative data in their pilot trials.
Beets 2022	Impact of risk of generalizability biases in adult obesity interventions: A meta-epidemiological review and meta-analysis	conduct	The purpose of this study was to build upon previous evidence of the influence of RGBs[risk of generalizability biases] and evaluate their impact in a sample of published pilot/feasibility studies and larger scale trials of the same behavioral intervention on a topic related to adult obesity.	The primary testing of the impact of the biases was performed by comparing the change in the SDM [standardized difference of means] from the pilot study to the larger scale trial for studies coded with and without a given bias present.	For pairs where the pilot was coded as having an RGB present and in the larger scale trial, the RGB was no longer present, the SDM decreased by an average of ΔSDM 0.41, range 1.06 to 0.01.
Beets 2020	Identification and evaluation of risk of generalizability biases in pilot versus efficacy/ effectiveness trials: a systematic review and meta-analysis	design	The purpose of this study was to describe the rationale for generating an initial set of “risk of generalizability biases” (defined below) that may lead to exaggerated early discoveries and therefore increase the risk of	The primary testing of the impact of the biases was performed by comparing the changing in the [standardized mean difference] SMD from the pilot study to the larger, efficacy/ effectiveness trial for	For four of the generalizability biases – delivery agent, implementation support, intervention duration, and measurement – the difference in the SMD (i.e., the larger, more well-powered trial SMD minus

			subsequent efficacy and effectiveness trials being unsuccessful.	studies coded with and without a given bias present.	the pilot SMD) was larger in the pairs of pilot studies that had the bias present and subsequently did not have the bias present in the larger, more well-powered trials, compared to pairs that did not have the biases present. Specifically, the change in the SMD was – 0.325 (95CI – 0.556 to – 0.094) for agent delivery, – 0.346 (– 0.640 to – 0.052) for implementation support, – 0.342 (– 0.498 to – 0.187) for intervention duration, and – 0.360 (– 0.631 to 0.089) for measurement.
Bhatt 2018	Evaluating Completeness of Reporting in Behavioral Interventions Pilot Trials	reporting	The current systematic survey aims to evaluate published pilot and feasibility trials of behavioral interventions in terms of quality of reporting using the newly published CONSORT pilot extension as a benchmark. More specifically, we aim to: 1. evaluate the quality of reporting as measured by completeness of reporting based on the 2016 CONSORT extension to pilot and feasibility trials among pilot and feasibility	The primary outcome of this survey was to evaluate the pro portion of reported items on the CONSORT statement extension for randomized pilot and feasibility trials.	The mean CONSORT reporting score across all included articles was 51.6% (SD ¼ 15.1).

			trials investigating behavioral interventions in clinical populations		
Charlesworth 2020	A bibliometric analysis of the conversion and reporting of pilot studies published in six anaesthesia journals	reporting	We hypothesised that only a small proportion of published articles in anaesthesia journals describing 'pilot' or 'feasibility' studies were correctly labelled, and similarly, that only a small proportion were converted into formal randomised, controlled trials. We decided to conduct a bibliometric review of pilot studies reported in anaesthesia journals to test this hypothesis.	The aim of this study was to determine the number of pilot studies correctly labelled as such and report on their characteristics as compared with included studies which were not in preparation for a larger trial.	A total of 34 (12.8%) studies were correctly labelled as pilot or feasibility studies in preparation for a larger trial.
El-Khechen 2021	Design, analysis, and reporting of pilot studies in HIV: a systematic review and methodological study	design, analysis, reporting	The primary goal of this study was to describe the design, analyses, and reporting of pilot studies in HIV.	The main outcomes of interest were the following: 1. The nomenclature of pilot studies ("pilot" or "feasibility" in the title)	Most studies were easily identifiable as pilot or feasibility studies, with 179 studies (72.2%) including the terms pilot, feasibility, or a feasibility outcome in the study title to denote the pilot status. The remaining 69 studies (27.8%) had no indication of their pilot nature in the title.
Horne 2018	Pilot trials in physical activity journals: a review of reporting and editorial policy	reporting	The overall aim of this work is to evaluate the impact of the CONSORT 2010 extension in the field of physical activity. This will be	A data extraction form was developed using the CONSORT 2010 extension as a guide.	Notable differences were found across the three categories of trials in terms of how they identified in the title of the article. FDT

			done by reviewing the reporting of pilot trials in physical activity journals before and after the 2016 publication of the CONSORT 2010 extension. This first article presents a review of articles published in 2012–15.		[future definitive trial] and FNI [feasibility of a novel intervention but did not explicitly address a FDT in their objectives] trials varied as to whether they identified as pilot or feasibility, all of them identifying as either or both. In contrast, 12 (71%) of the NF trials identified as pilot in the title, and none as feasibility. In addition, only one (6%) of the NF trials identified as randomised, compared with five (71%) and three (43%) of the FDT and FNI studies respectively.
Isiguzo 2018	Quality of pilot trial abstracts in heart failure is suboptimal: a systematic survey	reporting	The aims of this survey are to (1) evaluate the quality of reporting of abstracts of pilot RCTs in the past 26 years (1990–2016), using the CONSORT extension for reporting of abstracts of pilot trials;	We calculated the percentage of trials that scored yes on each of the 16 items and the associated 95% confidence interval (CI).	None of the studies reported all the 16 items in the checklist (Table 2), the maximum reported number of items was 12, with a mean (SD) of 8.3 (1.7) items.
Isiguzo 2017	Quality of abstracts of pilot trials in heart failure: A protocol for a systematic survey	reporting	To evaluate the reporting quality of abstracts of pilot trials in heart failure in the past 26 years (1990–2016), using CONSORT extension for reporting of abstracts of pilot trials as the reference standard.	Primary outcome measures: Reporting quality by presenting reported and unreported items on the checklist as frequencies and percentages using descriptive analysis.	N/A (protocol)

Kenis 2020	Do radiological research articles apply the term "pilot study" correctly? Systematic review	design, reporting	Therefore, the aim of the present systematic was to determine the proportion of radiological “pilot” studies that use this description correctly.	Not stated	Ultimately, no individual study qualified as a genuine pilot study when assessed against the a priori criteria… Notably, 66 (84.6%) studies presented measures of test accuracy and were framed as studies of diagnostic test accuracy. Twelve studies presented elements of feasibility, and eight elements of technology assessment.
Kosa 2019	Reporting quality of pilot clinical trials in chronic kidney disease patients on hemodialysis: a methodological survey	reporting	Using the newly published CONSORT extension for pilot trials, we undertook a methodological survey to assess reporting completeness among pilot trials investigating interventions in CKD patients on hemodialysis (HD patients) and explored factors associated with better completion of reporting.	The primary outcome of this survey was the completeness of reporting of each of the items on CONSORT statement extension for randomized pilot and feasibility trials checklist, measured as a number and proportion of studies reporting each of the 40 items.	The mean CONSORT reporting score across all included articles was 18.4 (SD = 4.4, minimum = 8, maximum = 29) out of a possible 34 items (6c, 7b, 11a, 11b, 18, and 19a were excluded as they had a not applicable option).
Kristunas 2019	The current use of feasibility studies in the assessment of feasibility for stepped-wedge cluster randomised trials: a systematic review	design	This review aims to gain an insight into how feasibility studies are being used to inform the design of SW-CRTs [stepped wedge cluster randomized trials]. Specifically, our objectives were to: Systematically identify published feasibility	Not stated	Of these 11 studies included, less than half reported that the study had been registered (Table 2) [Additional File 3]. The majority of the identified studies were reports of results, but three protocols were also identified. Just

			studies designed to inform SW-CRTs;		over half of the studies described themselves as a pilot study, with the others using terms such as “feasibility study”, “acceptability and feasibility pilot”, “consultation exercise” or “formative research”.
Lawson 2022	Pilot and Feasibility Studies in Rehabilitation Research: A Review and Educational Primer for the Physiatrist Researcher	conduct, reporting	This article has been structured from an educational perspective based on the following key aims: (1) to summarize current practices in the conduct and reporting of pilot and feasibility studies in rehabilitation research;	Not stated	Approximately a third (34.0%) of studies stated a primary objective related to feasibility (e.g., the goal of the study is “to examine (1) feasibility...,” “to establish if the treatment could be safely tolerated,” “to evaluate patient acceptance and satisfaction with the treatment...”).
Mailhot 2020	Methodological reporting in feasibility studies: a descriptive review of the nursing intervention research literature	design, reporting	Thus, our objective was threefold: (a) to systematically review the literature on feasibility studies in nursing intervention research;	Not stated	The majority of papers included the label ‘pilot’ or ‘feasibility’ in their title (n¼142, 76.3%); the remainder included ‘pilot’ or ‘feasibility’ in their abstract (n¼32, 17.2%) or main text (n¼12, 6.5%). In addition to the label ‘pilot’ or ‘feasibility,’ the title of 26 papers (14%) included labels suggesting formal hypothesis testing (e.g. ‘efficacy’, ‘effectiveness’, ‘effect’ or ‘impact’).

Mbuagbaw 2019	The reporting of progression criteria in protocols of pilot trials designed to assess the feasibility of main trials is insufficient: A meta-epidemiological study	reporting	Our objectives were to describe reporting of progression criteria to main trial and to determine the factors associated with reporting of progression criteria.	Not stated	The proportion of studies reporting progression criteria [45 (19.8%); 95% CI 14.8–25.6]
Mellor 2021	Progression from external pilot to definitive randomised controlled trial: a methodological review of progression criteria reporting	design, reporting	The primary objective was to describe the reporting of progression criteria, including the areas of feasibility that progression criteria were based on as described in a published framework of reasons for conducting pilot trials, their rationale or justification and who established and assessed the progression criteria.	Not stated	The reported progression criteria generally addressed some (99/160, 62%) or all (53/160, 33%) of the pilot trial's feasibility outcomes. The pilot trial publications reported a mean of 4 (mean 4.05) progression criteria targets per pilot trial. Recruitment (113/160, 71%) and retention (106/160, 66%) were the most commonly reported indicators of feasibility to inform progression. In total, we identified 58 distinct areas of trial feasibility that contributed to progression criteria, which we grouped into four domains: process, resource, management and scientific.

Rosala-Hallas 2019	A review of current practice in the design and assessment of internal pilots in UK NIHR clinical trials	design	The aim of this study was to provide an insight into current practice regarding internal pilots in clinical trials, specifically the design and review process, by assessing a cohort of studies funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme.	Not stated	Thirty-six trials (63%) specified the progression criteria in the latest available version of the protocol and of the 22 with available previous versions, 19 (86%) specified the progression criteria in all available previous versions of the protocol.
Scola 2019	A methodological survey on reporting of pilot and feasibility trials for physiotherapy interventions: a study protocol	reporting	The purpose of this methodological survey is to describe the quality of reporting of abstracts and full articles of pilot or feasibility trials from a representative sample in the field of physiotherapy. Specifically, the first aim is to determine the percentage of trial reports indexed in the Physiotherapy Evidence Database (PEDro), which claim to be pilot or feasibility trials that evaluate feasibility.	The primary analysis will be a descriptive analysis of completeness of reporting of the abstracts and full articles of the pilot or feasibility trials.	N/A (protocol)

CHAPTER 6: Discussion

6.1 Discussion

This section is structured first addressing the objectives of the studies in this thesis (part 1), and second, by providing a discussion of the implications of these findings (part 2).

Part 1:

Addressing the research objectives

The following is a summary of the key findings from this thesis:

- Researchers that carry out methodological studies of health research often do so with the intent to advise on improvements to current research practice (i.e., design, conduct, analysis, and reporting) in the form of methods guidance.
- Based on a case study of a sample of pilot and feasibility studies (PAFS) in rehabilitation research, the resulting methods guidance focused on addressing identified issues regarding overall design, improper naming of studies, lack of feasibility objectives, outcomes and progression criteria, and sample size justification.
- It is possible that a number of the identified studies (possibly close to half) labelled as PAFS were in fact small, underpowered studies that were likely not originally designed to measure the feasibility of definitive research. Although reporting guidelines and definitions for PAFS have only been published since 2016—and therefore this finding should be interpreted with caution—this provides a glimpse into the idea of research waste (i.e., "improper" study design, important outcomes not assessed, and patient contributions to research rendered unproductive).

- Based on the pilot study of methodological studies indexed in PubMed, there were many names used to describe studies with overall similar designs and objectives (i.e., syntheses of research practice based on published literature) including 'methodological review', 'systematic review', and 'systematic survey'.
- The diversity in names and reporting styles confirmed a lack of consensus in the research community on what these studies should be called, hampering efforts to identify them in literature databases. However, a preliminary search strategy was sensitive enough to detect methodological studies in PubMed, and overall, it was found that it would be feasible to progress to a full review of multiple databases.
- Methodological studies that evaluated PAFS were investigated in a full review of five databases and additional sources. The above findings were corroborated in terms of the diversity in names (e.g., 'review', 'survey', 'methodological review', variations of 'meta-epidemiology' and so on). The most common research practice investigated was the reporting of research.
- Existing reporting checklists were sometimes used, but authors only referred to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Completed checklists were never provided, possibly due to some incompatibility with methodological study design features.
- Overall, the findings in this thesis are suggestive of the importance of consensus on reporting and nomenclature for making methodological research more accessible and findable. Efforts to streamline the naming of methodological studies would be valuable to the health research community and end-users of methodological studies.

Methodological contributions

This thesis is distinct in that it is interrogating the field of methodological research itself by evaluating its reporting, methodology, and knowledge gaps in different ways. The studies comprising this thesis are characterized by several diverse designs. Chapter 2 is a methods guidance paper, commissioned by the journal for the purpose of educating clinicians and clinical researchers about the appropriate way to design, carry out, and report PAFS. Chapter 3 describes a pilot study with explicit progression criteria and includes pilot-testing to determine the sensitivity and specificity of a search strategy. Chapter 4 describes a reporting guideline development protocol. Chapter 5 describes a methodological review. Together, these works demonstrate a suite of competencies for methodological research practice.

Part 2:

Bias in methodological studies

The implications from bias in methodological research are less clear than that for primary clinical research, which more directly impacts patients and therefore, is likely easier to observe. Based on a survey of researchers that conduct methodological studies (those employing systematic review methods) by the Methodology Programme of the National Health Service Health Technology Assessment Programme in UK, researchers mentioned the need to reduce bias as a key recommendation (Lilford et al., 2001). Bias in methodological research is important to address since “many arguments depend on underlying philosophical assumptions or other perspectives” (Lilford et al., 2001). This is in line with previous investigations by researchers looking at publication bias in studies of

publication bias (Dubben & Beck-Bornholdt, 2005), i.e., the preferential publication of positive findings. They concluded that no publication bias was found, based on a funnel plot, although the power to detect bias was low due to the small sample size. However, others argue that methodological studies are equally prone to publication bias as other research (Song, 2005). It is good research practice to avoid or strive to reduce biases in all scientific inquiry, and it remains to be seen how studies of methodological studies might influence this goal.

Pre-registration of research and data-sharing can help to address issues of bias (Bradley et al., 2020; Mann, 2005; Sijtsma, Emons, Steneck, & Bouter, 2021), and platforms promoting research transparency like the Open Science Framework (Center for Open Science, 2012) are now regularly used by methodologists to pre-register their work. In the context of methodological research, this practice is largely voluntary. Other initiatives such as the Research on Research registry (School of Healthcare Enterprise and Innovation, 2023) have sprung up since for this purpose. Out of this need for transparency, the Evidence-Based Research (EBR) Network was created in 2014 with the purpose of advocating for an evidence-based approach to research (Evidence-Based Research Network, 2023). They advocate for the use of prior research in systematic and transparent ways to inform new studies, stating that research questions should be addressed in an "accessible manner" (Evidence-Based Research Network, 2023). Similarly, reporting guidelines are useful tools in promoting the transparency of research practice, although their usefulness is dependent on whether they are used. Buy-in from different stakeholders, such as journal editors, is key and such stakeholders have been

recruited as part of the development of MISTIC. However, researchers may still choose not to publish their work even if they plan to follow a suitable reporting guideline. Therefore, pre-registration should be a key reporting item in a future reporting guideline for methodological studies.

Intent of methodological studies

In Chapter 5, we have found that many authors of methodological studies intend to inform future research practices at least broadly, and in some cases, educate others in their field by highlighting issues. Earlier studies such as the work by Schulz (Schulz, Chalmers, Hayes, & Altman, 1995) investigated impacts on treatment effects in trials, for example. Referring to the previous example, where do methodological studies of publication bias in studies of publication bias fit in? Kranke argues that such studies, apart from illustrating that publication bias exists, are not meaningful for clinicians and patient care (Kranke, 2005). Scientists are free to pursue limitless research questions, and such efforts may be in the name of pure theoretical and academic pursuit. However, in addition to being producers of knowledge, researchers are faced with the reality that they are in fact "performing a public role" (Reydon, 2019). This comes with certain obligations since researchers have privileged access to specific areas of knowledge that the general public does not (Reydon, 2019). In line with this, it is the responsibility of the research community to, at the very least, make their research accessible and transparent. In so doing, this will help foster greater public trust in science. When studies (both clinical and methodological) with important downstream consequences for patients are made more accessible, the findings may be readily accepted by a wider community.

Allowing diverse audiences to access health research more readily can lead to a greater acceptance of and appreciation for the value of methodological research.

Future investigation might warrant a closer look at the composition of research teams and the reasons for conducting methodological work. Many authors report challenges faced when conducting clinical studies as a motive for their methodological studies. This makes sense since the methodology of clinical research and clinical research go hand-in-hand. A qualitative study of authors, including individuals conducting methodological work who are not trained as methodologists, can help to uncover motives for the generation of such a large body of research. For example, are researchers undertaking methodological studies due to a lack of methods guidance and resources on their research teams?

Reporting guidelines for better transparency and access to methodological studies

Reporting is a key component of research. Several authors have stated that the final published report of research is its only lasting record (Lang, 2010) and the only evidence that the research ever occurred (Rennie, 1986), although this has changed with online communication/media platforms and practices such as integrated knowledge translation. General formatting guidance for articles submitted to biomedical journals has been produced in 1978 by the International Committee of Medical Journal Editors, and has since evolved beyond formatting recommendations (International Committee of Medical Journal Editors, 2023). Based on a Google Scholar search in July 2023, the tutorial for conducting methodological studies published by the lead MISTIC investigators (Mbuagbaw, Lawson, Puljak, Allison, & Thabane, 2020) has been cited 54

times by non-authors/non-collaborators of the project since its publication in 2020. This suggests an ongoing appetite for methodological research, and consequently, this is expected to continue to present challenges with its classification. Similarly, the MISTIC protocol, published in 2020, has been cited 23 times by non-authors/non-collaborators. Most often, researchers mentioned that there is no relevant guideline for their study design as a reason. Alternative guidelines for reporting methodological studies have been proposed (Martin et al., 2020; Murad & Wang, 2017) and in some instances, cited by researchers to report their work (Puljak, Makaric, Buljan, & Pieper, 2020). Although there are various acceptable approaches to development, none of the previous guides have been produced through a formal reporting guideline development process which is advocated by experts (Moher, Schulz, Simera, & Altman, 2010).

Like most research, reporting guidelines are not without criticism. Some experts have suggested expiry dates for reporting guidelines to keep them relevant since the methodology of health research constantly evolves (Rothman & Poole, 2007). Others fear that they stifle research creativity, or that "variant designs might be considered as not conforming" and therefore suggestive of lesser quality (Vandenbroucke, 2009). Health research reporting guidelines were not developed with the intent to assess the quality of studies (Logullo, MacCarthy, Kirtley, & Collins, 2020) although they continue to be wrongfully used for such purposes. While some are understandably resistant to the idea, others have critiqued reporting guidelines on the basis that they are "condescending" and that researchers including peer reviewers and editors are already experts not in need of checklists to perform their tasks (MacMahon & Weiss, 2007). Such reasoning begs the

question, should researchers investigating reporting in other studies for example, be exempt from transparency in reporting themselves? Unfortunately, researchers are also incentivized to publish on the basis of quantity and appearance, not quality or correctness, which brings about other issues (e.g., incorrect research is rarely highlighted, author replies to letters that highlight issues can be dismissive) (Van Calster, Wynants, Riley, van Smeden, & Collins, 2021).

Methodological research is often done by researchers with specialized training in research methods. Editors and peer reviewers have varied expertise that does not necessarily include research methodology, especially in medical journals targeting clinical audiences. It is also known that there is a deeply rooted history of academic elitism in higher education, and this attitude does not translate well to the accessibility of research (which includes transparency). This is especially true for those without specialized training and for those who the majority of health methods research intends to impact downstream: patients. Interpretations of reporting guidelines as condescending or as prescriptive are possibly missing the point: that they are meant to be helpful tools and reminders to promote clarity. This may depend on journals, some of which require the submission of completed checklists with manuscripts. On a positive note, reporting guidelines have been lauded as useful educational tools by students (Vandenbroucke, 2009). Similarly, a reporting guideline for methodological studies may increase trainees' awareness of this field, for which courses have been reported as still rare (Weissgerber, 2021).

Researchers might argue that reporting cannot and should not be standardized in this field. There are likely various nuances in the types of methodological studies across different fields of health research (e.g., biostatistics studies compared to diagnostic studies). It is important to highlight the heterogeneity among methodological studies, and that the intent of streamlining reporting and nomenclature does not equate to the prescription of conduct of the work itself, or that it should be pigeon-holed. As a result, there may be alternative approaches to reporting methodological studies that will need to be considered in the reporting guideline development process. For example, a multi-dimensional reporting guide directing authors to existing guidelines (e.g., PRISMA), as needed, may receive greater acceptance.

As has been presented in this thesis, the variety of naming conventions compromise the use of and access to methodological research, and our belief is that this is the key issue in this field. Proponents of EBR have stated that "All researchers should be able to search for, critically appraise, and interpret systematic reviews in the context of new study results" (Lund et al., 2016). A similar principle should be applied to methodological studies of health research. The work presented in this thesis also supports the need for a dedicated database to enable easy access, or other approaches such as the development of appropriate MeSH terms. Balancing opposing views on this topic will present many challenges and continued discussion between end-users is essential for this purpose. A diverse array of stakeholders in the guideline development process, user acceptability and testing initiatives, and broad knowledge translation activities will be essential to this endeavour.

Conclusions

This thesis describes one of the first set of studies to broadly characterize and synthesize information about methodological studies in the health research field. Methodological studies represent a rich body of research, characterized by diverse names, methodologies, and reporting styles. It was important to assess the earliest documented reports as well as contemporary practices. We aimed to produce an exhaustive and comprehensive overview of these studies to the extent possible, and in this thesis, the focus was narrowed to methodological studies of PAFS as a case study. The results of the full review that is underway, described in Chapter 4, suggests promise in the ability to characterize this field more fully. However, this field is marked by a heterogeneous body of literature and buy-in from researchers and stakeholders, including journal editors, will be essential to success for streamlining the reporting of and improving access to methodological studies.

6.2 References

Bradley, S. H., DeVito, N. J., Lloyd, K. E., Richards, G. C., Rombey, T., Wayant, C., & Gill, P. J. (2020). Reducing bias and improving transparency in medical research: a critical overview of the problems, progress and suggested next steps. *J R Soc Med, 113*(11), 433-443. doi:10.1177/0141076820956799

Center for Open Science. (2012). Open Science Framework (OSF). Retrieved from https://osf.io

Dubben, H. H., & Beck-Bornholdt, H. P. (2005). Systematic review of publication bias in studies on publication bias. *BMJ, 331*(7514), 433-434. doi:10.1136/bmj.38478.497164.F7

Evidence-Based Research Network. (2023). About Evidence-Based Research (EBR). Retrieved from http://ebrnetwork.org/about-ebr/

International Committee of Medical Journal Editors. (2023). History of the Recommendations. Retrieved from https://www.icmje.org/about-icmje/contact-icmje/

Kranke, P. (2005). Review of publication bias in studies on publication bias: meta-research on publication bias does not help transfer research results to patient care. *BMJ, 331*(7517), 638. doi:10.1136/bmj.331.7517.638

Lang, T. (2010). CONSORTing With a QUOROM of MOOSES: The Standards Movement in Scientific Reporting. Neurourology and Urodynamics, 29, 28-29. doi:10.1002/nau

Lilford, R. J., Richardson, A., Stevens, A., Fitzpatrick, R., Edwards, S., Rock, F., & Hutton, J. L. (2001). Issues in methodological research: perspectives from researchers and commissioners. *Health Technol Assess, 5*(8), 1-57. doi:10.3310/hta5080

Logullo, P., MacCarthy, A., Kirtley, S., & Collins, G. S. (2020). Reporting guideline checklists are not quality evaluation forms: they are guidance for writing. Health Sci Rep, 3(2), e165. doi:10.1002/hsr2.165

Lund, H., Brunnhuber, K., Juhl, C., Robinson, K., Leenaars, M., Dorch, B. F., . . . Chalmers, I. (2016). Towards evidence based research. *BMJ, 355*, i5440. doi:10.1136/bmj.i5440

MacMahon, B., & Weiss, N. S. (2007). Is there a dark phase of this STROBE? *Epidemiology, 18*(6), 791. doi:10.1097/EDE.0b013e3181570f6e

Mann, J. (2005). Review of publication bias in studies on publication bias: mandatory publication of data may help. *BMJ, 331*(7517), 638. doi:10.1136/bmj.331.7517.638-b

Martin, G. P., Jenkins, D. A., Bull, L., Sisk, R., Lin, L., Hulme, W., . . . Predictive Healthcare Analytics, G. (2020). Toward a framework for the design, implementation, and reporting of methodology scoping reviews. *J Clin Epidemiol, 127*, 191-197. doi:10.1016/j.jclinepi.2020.07.014

Mbuagbaw, L., Lawson, D. O., Puljak, L., Allison, D. B., & Thabane, L. (2020). A tutorial on methodological studies: the what, when, how and why. *BMC Med Res Methodol, 20*(1), 226. doi:10.1186/s12874-020-01107-7

Moher, D., Schulz, K. F., Simera, I., & Altman, D. G. (2010). Guidance for developers of health research reporting guidelines. *PLoS Med, 7*(2), e1000217. doi:10.1371/journal.pmed.1000217

Murad, M. H., & Wang, Z. (2017). Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med, 22*(4), 139-142. doi:10.1136/ebmed-2017-110713

Puljak, L., Makaric, Z. L., Buljan, I., & Pieper, D. (2020). What is a meta-epidemiological study? Analysis of published literature indicated heterogeneous study designs and definitions. *J Comp Eff Res, 9*(7), 497-508. doi:10.2217/cer-2019-0201

Rennie, D. (1986). Guarding the guardians: a conference on editorial peer review. JAMA, 256(17), 2391-2392.

Reydon, T. A. C. (2019). *What attitude should scientists have? Good academic practice as a precondition for the production of knowledge* (1st ed.). New York, NY: Routledge.

Rothman, K. J., & Poole, C. (2007). Some guidelines on guidelines: they should come with expiration dates. Epidemiology, 18(6), 794-796. doi:10.1097/EDE.0b013e3181571259

School of Healthcare Enterprise and Innovation, N. E., Trials and Studies Coordinating Centre. (2023). Research on Research registry and community hub. Retrieved from https://ror-hub.org/

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA, 273*(5), 408-412. doi:10.1001/jama.273.5.408

Sijtsma, K., Emons, W. H. M., Steneck, N. H., & Bouter, L. M. (2021). Steps toward preregistration of research on research integrity. *Res Integr Peer Rev, 6*(1), 5. doi:10.1186/s41073-021-00108-4

Song, F. (2005). Review of publication bias in studies on publication bias: studies on publication bias are probably susceptible to the bias they study. *BMJ, 331*(7517), 637-638. doi:10.1136/bmj.331.7517.637-c

Van Calster, B., Wynants, L., Riley, R. D., van Smeden, M., & Collins, G. S. (2021). Methodology over metrics: current scientific standards are a disservice to patients and society. J Clin Epidemiol, 138, 219-226. doi:10.1016/j.jclinepi.2021.05.018

Vandenbroucke, J. P. (2009). STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUOROM, REMARK... and CONSORT: for whom does the guideline toll? J Clin Epidemiol, 62(6), 594-596. doi:10.1016/j.jclinepi.2008.12.003

Weissgerber, T. L. (2021). Training early career researchers to use meta-research to improve science: A participant-guided "learn by doing" approach. PLoS Biol, 19(2), e3001073. doi:10.1371/journal.pbio.3001073