**Open Science interventions to improve reproducibility and replicability of research: a scoping review**

Leonie Dudda1,2, Eva Kormann3, Magdalena Kozula4, Nicholas J. DeVito5, Thomas Klebel10, Ayu P.M. Dewi6, René Spijker7,8, Inge Stegeman1,2, Veerle Van den Eynden9, Tony Ross-Hellauer3,10, Mariska M.G. Leeflang\*6

1. Department of Otorhinolaryngology and Head & Neck Surgery University Medical Center Utrecht, Utrecht, The Netherlands.
2. Brain Center, University Medical Center Utrecht, Utrecht, The Netherlands.
3. Institute of Interactive Systems and Data Science, Graz University of Technology, Graz, Austria.
4. Faculty of Psychology and Educational Sciences, Methodology of Educational Sciences Research Group, KU Leuven, Leuven, Belgium
5. Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK
6. Epidemiology and Data Science, Amsterdam Public Health, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
7. Cochrane Netherlands, UMC Utrecht, Utrecht University, Utrecht, The Netherlands
8. Medical Library, Amsterdam Public Health, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
9. Research Data Management Competence Centre, KU Leuven, Leuven, Belgium
10. Open & Reproducible Research Group, Know-Center GmbH, Graz, Austria

\* M.M.G. Leeflang is corresponding author.

**Abstract**

Various interventions – especially those related to open science – have been proposed to improve the reproducibility and replicability of scientific research. To assess whether and which interventions have been formally tested for their effectiveness in improving reproducibility and replicability, we conducted a scoping review of the literature on interventions to improve reproducibility. We systematically searched [Medline](https://www.medline.com/), [Embase](https://www.embase.com/), [Web of Science](https://www.webofscience.com/wos/woscc/basic-search), [PsycINFO](https://www.apa.org/pubs/databases/psycinfo), [Scopus](https://www.scopus.com/home.uri) and [Eric](https://eric.ed.gov/), on August 18, 2023. Grey literature was requested from experts in the fields of reproducibility and open science. Any study empirically evaluating the effectiveness of interventions aimed at improving the reproducibility or replicability of scientific methods and findings was included. An intervention could be any action taken by either individual researchers or scientific institutions (e.g., research institutes, publishers and funders). We summarized the retrieved evidence narratively and in an evidence gap map. Of the 104 distinct studies we included, 15 directly measured the effect of an intervention on reproducibility or replicability, while the other research questions addressed a proxy outcome that might be expected to increase reproducibility or replicability, such as data sharing, methods transparency or preregistration. Thirty research questions within included studies were non-comparative and 27 were comparative but cross-sectional, precluding any causal inference. Possible limitations of our review may be the search and selection strategy, which was done by a large team including researchers from different disciplines and different expertise levels. Despite studies investigating a range of interventions and addressing various outcomes, our findings indicate that in general the evidence-base for which various interventions to improve reproducibility of research remains remarkably limited in many respects. The study protocol was pre-registered on the Open Science Framework (OSF) under (DOI 10.17605/OSF.IO/D65YS. This work was funded by the European Union’s Horizon Europe research and innovation program under grant agreements No. 101094725 (OSIRIS) and 101094817 (TIER2).

**Introduction**

The robustness and trustworthiness of research results are in question [1–3]. This is true especially with respect to their *reproducibility* (defined in this paper as obtaining the same or similar results when rerunning analyses from previous studies using the original design and data and code [cf., 4]) and *replicability* (defined here as obtaining the same or similar results when repeating, in whole or part, a prior study [cf., 4]). Ensuring that the results of studies are robust and can be independently confirmed, is expected to reduce waste [2,5] and lead to more reliable outcomes that better inform evidence-based decisions [6,7]. Furthermore, studies that can be independently confirmed may increase public trust in the scientific enterprise [8,9]. Reproducibility and replicability therefore underpin the credibility and reliability of research findings in many research fields, especially in science, technology, engineering and mathematics. From now on in the manuscript, for brevity, we will use the word “reproducibility” to refer to both terms, as well as the reproducibility of methods or code, etc. (i.e., what Goodman et al. [10] call “methods reproducibility”). Where more narrow usage of terms, per Nosek et al. [4] is intended, we will state this explicitly.

Debates on reproducibility have gained increasing prominence in various scholarly fields, as well as in the general press [11,12]. Pivotal to this debate were failures to reproduce study findings across the medical, behavioural and social sciences. Fields like psychology [13], biomedical research [14,15], economics [16], and broader social sciences [17] all witnessed “many-lab" studies whose findings indicated levels of reproducibility ranging between 30 and 70%. A piece in Nature News in 2016 reported survey findings (ironically themselves lacking in rigour and transparency) that highlighted that between 60% and 80% of scientists across various disciplines encountered hurdles in reproducing the work of their peers, with similarly noteworthy difficulties encountered when attempting to replicate their own experiments (40% to 60%) [1]. Other scholars have argued that the meanings and relevance of “reproducibility” vary widely across research fields and argue the unsuitability of these concepts – at least as strictly as it is understood in more experimental domains – as a criterion of quality for within humanities [29–31] or qualitative research [32][33–35].

Given these interdisciplinary differences, the applicability of reproducibility-related interventions (what works, in which circumstances) must be expected to vary. Ongoing discussions highlight various strategies and practices that have the potential to address the rigour and reliability of large areas of scholarly work through improving the quality of methods, reporting and verification of research [19]. Factors that have been associated with perceived poor levels of reproducibility, include selective non-publication, questionable research practices, insufficient training in research methods and a lack of transparency and data accessibility [5,18–20]. Although amongst proponents of Open Science, the consensus is that openness of methods, materials and community will improve reproducibility of science [24], inadequate access to the necessary data for rerunning experiments or analyses remains one of the major concerns in science [1,21,22] [23].

Although multifaceted efforts underscore a shared commitment towards reproducibility within much of the academic landscape, the extent to which practices and interventions have been empirically investigated for their effectiveness to improve reproducibility, across research disciplines and time, remains unclear [44]. To assess which interventions have been formally tested for their effectiveness in improving the reproducibility of science, we conducted a scoping review of the literature on this. By summarizing current evidence and highlighting gaps in knowledge, this study aims to provide insights for researchers, policymakers, and stakeholders invested in promoting reproducibility.

**Objectives**

The objective of this scoping review is to provide an overview of interventions that have been formally investigated for their effect on reproducibility and replicability in science. An intervention could be any action exhibited by either individual researchers or scientific institutions (e.g., research institutes, publishers and funders) with the aim to enhance reproducibility. We specifically focused on Open Science practices but interpreted their definition broadly.

**Methods**

**Study registration**

The study protocol was pre-registered on the Open Science Framework (OSF) (DOI [[10.17605/OSF.IO/D65YS](http://dx.doi.org/10.17605/OSF.IO/D65YS))](http://dx.doi.org/10.17605/OSF.IO/D65YS) and published on Open Research Europe (ORE) [45]. Deviations from the protocol are reported transparently in Supplementary File 1. All materials, including the data extraction sheets, data and code are available on OSF (DOI [[10.17605/OSF.IO/7EF5H](http://dx.doi.org/10.17605/OSF.IO/7EF5H))](http://dx.doi.org/10.17605/OSF.IO/7EF5H).

**Design: Scoping review**

Scoping reviews, also sometimes referred to as ‘mapping reviews’ or ‘scoping studies’ refer to an approach to evidence synthesis that are particularly helpful when the literature is complex and heterogeneous. They usually address broader review questions than traditionally more specific systematic reviews. As the notion of reproducibility and replicability varies across disciplines, we gathered a multidisciplinary team of authors and based the terms used throughout the scoping review on broad definitions taken from multiple texts and disciplines. We conducted the scoping review following the guidance provided by the Joanna Briggs Institute and we reported the scoping review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses: extension for Scoping Reviews (PRISMA-ScR) guidelines (see Supplementary File 6) [46].

**Eligibility criteria**

**Study designs**

We included studies evaluating the effectiveness of interventions aimed at enhancing the reproducibility or replicability of scientific methods and findings in a comparative design. These could be either between-subject or within-subject comparisons. Non-comparative studies were only included if the introduction of an intervention was explicitly stated (e.g. post-intervention study reporting the level of reproducibility after the implementation of a guideline). Studies that only report the prevalence of an intervention (e.g., measuring only the amount of preregistration in a field) were not included.

We excluded studies if they did not test an intervention; if the full text was unavailable; and if the article did not report primary resources, such as Stage 1 registered reports or position papers. We also excluded studies that investigated agreement in test results between observers or laboratories, which is also sometimes referred to as reproducibility of test results.

**Article Types**

We included any article type containing primary data, including peer reviewed original articles, early access papers and preprints. Position papers, study protocols, editorials and any other articles lacking primary data were excluded. Reviews were also excluded from our final synthesis, but their reference lists were checked for additional relevant publications. We used Zotero reference management software to check for any retracted studies within our corpus of articles (none were found).

**Participants**

Participants in eligible studies could be researchers of any career level, holders of other academic posts such as editors or administrative staff, or an entire academic institution such as a university, journal, publisher or funder. The included studies could cover any scientific field or discipline.

**Interventions and Comparators**

We included any intervention that aimed to improve the reproducibility and replicability of science, basing the categories of interventions on broad definitions taken from multiple disciplines. A list of interventions that we expected to include are listed in the protocol [45] and in Supplementary File 4. We included studies with any comparator.

**Outcomes**

Our primary outcome is the reproducibility or replicability of research. This means that studies attempt to replicate or reproduce another study, including aspects such as data collection, analyses, and/or conclusions (Supplementary File 4).

Since studies directly examining replicability or reproducibility were expected to be rare, we also defined a number of proxy outcomes that would be necessary, but not necessarily sufficient, for improved reproducibility. These proxies - practices commonly claimed within the literature to support reproducibility - were defined using a conceptual framework drawn from key texts [19,22,23,47], along with consultation experts on the measurement and assessment of reproducibility from different disciplines. Based on these considerations, outcomes such as preregistration, data sharing and completeness of reporting were therefore considered to be within the scope of the review as proxies that could facilitate reproducibility. As a consequence, various Open Science practices were included both as interventions and as proxy outcomes for reproducibility.

**Search strategies**

***Development of the search query***

The development of the search strategy was an iterative process, containing multiple steps and feedback loops. As the notion of reproducibility and replicability varies across disciplines, we based the search terms for interventions and outcomes on broad definitions taken from multiple texts and we searched ten databases to cover a broad variety in disciplines.

First, we requested “key articles” based on the inclusion criteria from the study team and a broader group of experts to develop and validate the search queries. This process initially identified five key articles. These were then subjected to citation coupling using the Connected Papers tool (https://www.connectedpapers.com/), which led to the identification of an additional eight key articles, ensuring a robust set of key articles. Throughout this phase, the number of articles was progressively reduced as we refined and discussed the inclusion criteria. Ultimately, only interventional articles were included to establish a definitive set of key articles. The intent was to base the search on articles that were unequivocally within scope, rather than those that were only partially relevant. This approach underscores the complexity of the project and highlights the challenge of defining clear boundaries within research definitions. The 15 key articles are available in a public Zotero folder (https://www.zotero.org/groups/4983756/osirisconsortium/collections/5Q3IJKZ2).

***Electronic search***

The electronic searches were done on August 18, 2023. We systematically searched [Medline](https://www.medline.com/), [Embase](https://www.embase.com/), [Web of Science](https://www.webofscience.com/wos/woscc/basic-search), [PsycINFO](https://www.apa.org/pubs/databases/psycinfo), [Scopus](https://www.scopus.com/home.uri) and [Eric](https://eric.ed.gov/). All databases were searched from inception until August 18, 2023. The search string was initially developed for Medline and then translated to the other databases. An initial search was constructed to align with our research group's understanding of the research question, yielding 16,930 references for screening. Following this initial screening, it was evident that certain topics were inadequately covered. Minor adjustments to the search strategy were implemented, resulting in an additional 472 references. After screening this initial batch, we discussed the outcomes and determined that the search did not adequately capture all areas of the research question. Consequently, a more sensitive search was developed to encompass the underrepresented topics from the initial results, aiming to achieve comprehensive coverage of the research question. This final search, while largely overlapping with the initial searches, introduced 18,661 additional unique references for screening leading to a total of 36,063 items for title and abstract screening. The final search strings are available in Supplementary File 2.

***Other searches***

We undertook an independent search for grey literature by reaching out to colleagues in October and November 2023, with e-mail requests for any relevant non-academic publications that address reproducibility interventions.

**Selection process**

All search results were collected in [EndNote](https://endnote.com/?language=en) 20 and duplicate references were removed using the software’s deduplication tool (http://dedupendnote.nl/). Any remaining duplicated were excluded manually during screening in EPPI Reviewer. At least two authors independently screened the reference lists of relevant reviews for relevant studies that were not discovered in other searches. The records identified with our search strategy were first screened for eligibility based on the title and abstract and then by reading the full manuscript. The exact process is described in further detail below.

***Title and abstract screening***

The title and abstract screening was done by L.D., E.K., M.K., N.J.D,, T.R-H., V.V.d.E., and M.L using Eppi Reviewer software. Initially, we planned to do the title and abstract screening in duplicate. However, due to the large volume of articles we opted for a single screening approach. Each assessor marked each title and abstract as *Include*, *Exclude*, or *Maybe*. The articles that were marked as *Maybe*, were then double-checked by one of the senior authors.

**Full text screening**

For articles that passed title and abstract screening, the full text was retrieved and reviewed by at least two members of the study team independently to determine final eligibility for the review. Full-text screening was conducted by authors L.D., E.K., M.K., N.J.D, T.K., A.P.M.D, V.V.d.E, T.R.H., M.L and additional contributors (see Acknowledgements). Disagreements during screening were solved by the senior authors T.R-H., M.L. and V.V.d.E, with additional discussion with the broader study team as needed.

To reduce the differences in screening between the authors, we started with a pilot screening of 23 randomly selected articles that were reviewed by the entire study team and then discussed. As new reviewers were added to the project, they completed a modified version of this pilot screening focusing on a selection of 10 articles from the original 23. If the manuscript of an article was in a foreign language, we either reallocated the article to a team member who spoke that language or used translation tools (Google Translate (<https://translate.google.com>) or DeepL translator (https://www.deepl.com/en/translator)) on the full text.

**Data-extraction**

A draft data extraction sheet, including extraction instructions and terminology definitions, was developed by L.D., M.K. and T.R-H. and reviewed by the whole team. Data extraction was done by L.D., E.K., M.K., N.J.D, T.K., A.P.M.D, V.V.d.E, T.R.H., M.L and additional contributors (see Acknowledgements). All extracting authors tested the extraction sheet using 10 articles and revised the final sheets based on pilot feedback. All fields extracted from included articles are detailed in Supplementary File 3. Data extraction was not conducted in duplicate as outlined in our protocol, primarily due to time constraints, but they were double checked another author. The data was extracted from the study manuscripts in EPPI reviewer. Because of time constraints we did not contact authors in case of insufficient data. If an article addressed multiple research questions referring to different outcomes, we separately extracted the data for each question. We therefore use the term “article” for the actual publication, and “study” for the separate outcomes.

**Data synthesis and analysis**

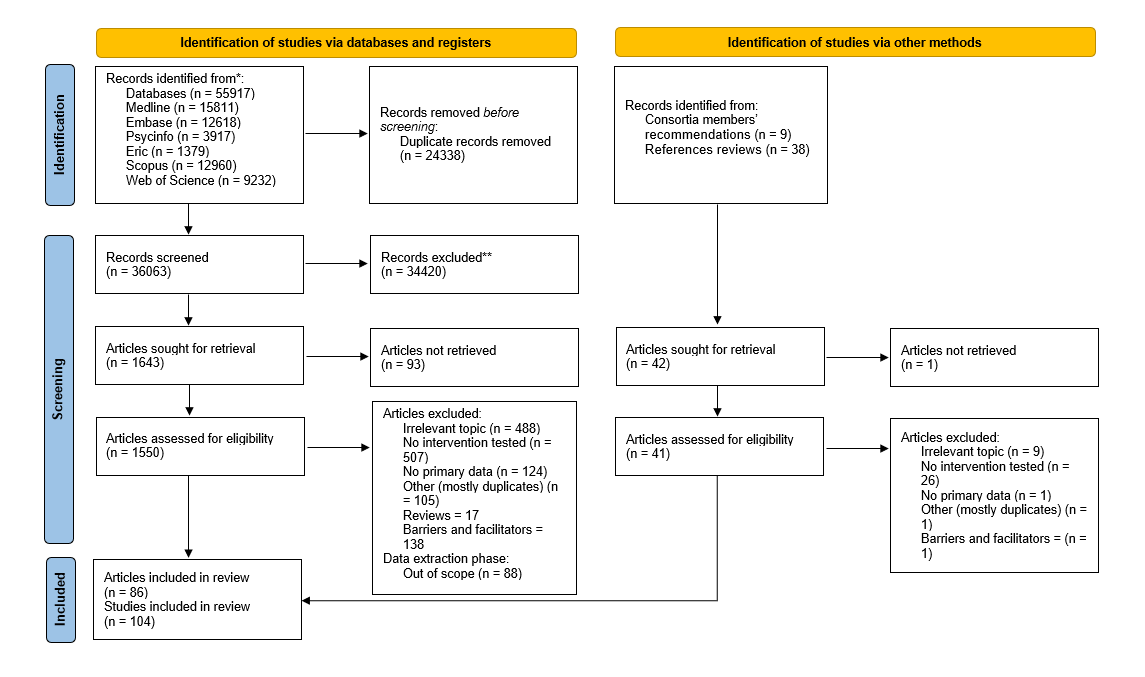
The data collected through EPPI reviewer were transferred to and cleaned in Microsoft Excel. We also used Microsoft Excel to summarize the characteristics of the included studies according to the interventions applied, the outcomes measured, the study design, and the stakeholders involved. The data extracted from the articles were presented graphically in evidence gap maps, which are a form of bubble plot with the different outcomes on the x-axis and the interventions on the y-axis. The size of the bubbles was defined by the number of studies that address this particular intervention/outcome combination and the colour of the bubbles was defined by the study design. If studies with different designs addressed similar intervention-outcome combinations, this was depicted using multiple colours. Information about stakeholders and academic field was represented using bar charts. We added extra variables to our dataset to enable more detailed analyses then were possible with the initial data-extraction items and we renamed variables to facilitate labelling in graphs and tables (Supplementary Files 3-5). All graphs were made using ggplot2 in R.

**Results**

Initially, we retrieved 36,063 articles; 1,643 of these passed the title and abstract screening. Following full-text screening, 172 articles were included for data-extraction. From the reference lists of relevant reviews, we assessed the full texts of additional 42 articles, of which we included three for data-extraction. During data-extraction, 88 further articles were excluded, because they were deemed out of scope. Most of these did not test a formal intervention (n=27) or assessed an irrelevant outcome (n=24), such as the power of a study or whether a guideline could be retrieved.

We thus included data from 86 reports in this scoping review (Figure 1) [48–133]. Most of these (n=70) assessed one intervention and one outcome, however 16 presented data for two (n=14) or three (n=2) separate hypotheses, separate research questions, separate sub-studies, or separate specific interventions or outcomes. Therefore, our analyses were based on 104 studies from 86 reports. Of these 104 studies, only 15 directly assessed the effect of one or more interventions on reproducibility or replicability by re-analysing data or actually reproducing the results of previous research. The remaining 89 studies investigated the effect of one or more interventions on a proxy outcome, such as data sharing or complete reporting. A glossary of definitions can be found in Supplementary file 6.

**Figure 1. PRISMA Flowchart of the search and selection process.**



***Characteristics of Included studies***

Table 1 shows the characteristics of included studies. Of the 86 included articles, 85 (99%) were peer reviewed, we included one preprint and one editorial about the implementation of registered reports by the editorial board. The included studies used a variety of study designs, as described below.

Randomized controlled trial (n=6 studies): five studies randomized subjects to receive the intervention or not [57,64,87,111,124]. One study randomized six sections within a manuscript and then authors randomly received instructions for three out of six sections, or for none of the sections in the control group [55].

Comparative between-subject design, cross-sectional (n=27 studies): These studies compared at one point in time (or during a fixed timeframe) participants exposed to a certain intervention with participants who have not been exposed to this intervention [60,75,81,85,92,96–98,101,102,104,109,115,120,121,123,127–130,132,133]. In four studies [60,85,104,123], the intervention was applied by researchers themselves. In five other studies, the intervention was not directly applied by the study team, as it consisted of a generic policy brief, or a publication. In all other cross-sectional studies, a journal, publisher or funder applied the intervention.

Comparative between-subject design, longitudinal (n=11 studies): These eleven studies measured the outcome at multiple different timepoints in different subjects (often articles), for example the measurement of reporting in articles over multiple years after the introduction of reporting policy in a journal [52,73,74,86,108,110,112,126]. One study followed a group of researchers who built a collaborative training community [110].

Comparative between-subject design, pre-post intervention (n=21 studies): This design was often applied in studies that checked reporting status of studies or materials- and methods-sharing status of studies before and after a certain event, for instance a change in journal policy [48,49,56,58,59,63,70,72,84,88,90,93,94,100,105,113,131].

Comparative between-subject design, non-randomized experimental (n=1): One study [125] experimented with a tool to help authors to adhere to writing guidelines. The authors could choose themselves whether they would use the tool or not.

Comparative within-subject design (n=7 studies): These seven studies issued a variety of designs, from measuring behavioural changes after a workshop [68] and standardization within one institute [118], to re-analyses of statistical findings [78,79 with two studies,128]. One study [71] did not report any quantitative results, instead the authors described the experiences after the implementation of an intervention. In all seven, the intervention was intentionally applied in an experimental setting.

Observational post-intervention only design (n=24 studies): This design was often applied in studies that checked reporting status of studies or materials- and methods-sharing status of studies after e.g. the impact of journal policy on these areas. There was no baseline measurement reported from before the implementation of the intervention [50,53,54,56,62,65–67,69,76,77,81–83,99,103,106,107,114,116,117,119].

Demonstration of a tool (n=6 studies): these are studies that apply a certain tool, technique, statistical solution or workflow to one or more use cases [51,61,89,91,95,122]. These studies are often limited to one project team, applying the intervention to themselves.

Other: there was one individual participant meta-analysis [80], which investigated associations between data availability statements and actual data availability, and associations with several other factors (e.g., journal policy, type of data, trial design, and human participants).

The unit of analysis for 77 of the 104 included studies was a published article (of which 13 were limited only to clinical trials). In these 77 studies, the number of articles investigated ranged from 3 to 2,121,580. Other units of analysis included were abstracts, analyses, datasets, institutions and individual researchers.

**Table 1. Characteristics of included studies (n=104).**

|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristic** | **Category** | **n** | **%** |
| Publication type | Peer-reviewed | 102 | 98.07 |
| Grey literature | 1 | 0.96 |
| Preprint | 1 | 0.96 |
| Study design | Randomized experiment | 5 | 4.81 |
| Comparative (between-subject, non-randomized) | 61 | 58.65 |
| Comparative (within-subject) | 7 | 6.73 |
| Post-intervention only | 25 | 24.04 |
| Demonstration of a tool | 6 | 5.77 |
| Other – individual patient meta-analysis | 1 | 0.96 |
| Academic discipline | Medical and health sciences | 57 | 54.81 |
| Social sciences | 22 | 21.15 |
| Natural sciences | 5 | 4.81 |
| Medical and health sciences & Social sciences | 6 | 5.77 |
| Medical and health sciences & Natural sciences | 8 | 7.69 |
| Social sciences & Natural sciences | 1 | 0.96 |
| Multidisciplinary | 5 | 4.81 |
| Intervention type | Open methodology | 10 | 9.62 |
| Rewards and incentives | 9 | 8.65 |
| Reporting guidelines, reporting standards | 15 | 14.42 |
| Policy guidelines (e.g. of funders and journals) | 49 | 47.12 |
| Open data and materials | 2 | 1.92 |
| Training | 2 | 1.92 |
| Open science tools | 7 | 6.73 |
| Other | 4 | 3.85 |
| Intervention implementer | Journal or publisher | 63 | 60.58 |
| Academic staff | 16 | 15.38 |
| Institution | 4 | 3.85 |
| Peer reviewer | 2 | 1.92 |
| Government | 1 | 0.96 |
| Multiple or other entities | 6 | 5.77 |
| Not applicable | 12 | 11.54 |
| Outcome type | Reproducibility | 9 | 8.62 |
| Replicability | 2 | 1.92 |
| Other direct outcomes | 4 | 3.85 |
| Registration status | 4 | 3.85 |
| Methods transparency | 35 | 33.65 |
| Research material sharing | 34 | 32.69 |
| Other | 16 | 15.38 |
| Author’s conclusion | Generally positive | 60 | 57.69 |
| Null / neutral | 43 | 41.35 |
| Generally negative | 1 | 0.96 |

***Interventions***

The grouping of interventions was difficult due to nuanced differences in apparently similar topics. For example, some studies examining reporting guidelines investigated whether these guidelines were mentioned in the author guidelines, while others examined whether journals actively requested completed checklists. Similarly, data sharing studies could focus on data sharing statements or actual data availability in a repository.

Some interventions could be seen as an exposure instead. For example, a board of editors issuing a statement about data sharing does not directly intervene in the research process but can still influence behaviour. This type of “exposure” could have – in principle – been subjected to randomized evaluations (e.g. journals randomly adopting a policy or not). Other interventions in our list, such as retractions or published guidelines cannot be randomized. For the purpose of this scoping review, we considered all these factors ‘interventions'. We grouped the interventions as explained below (Table 2).

Open methodology (n=10): Eight studies investigating some sort of preregistration of research plans. Five studies investigated the effect of preregistration [117, 127 with 3 studies, 129], one the effect of publishing the protocol of a study [75] and two studies investigated registered reports [106, 123]. Two other studies investigated the effect of open workflows [51, 118].

Policy guidelines issued by publishers and journals (n=46): The intervention most commonly studied was the implementation of a policy by publishers and journals. Seven of those investigated the effect of statements issued by the International Committee of Medical Journal Editors (ICMJE) on data sharing [56, 67, 114, 119], trial registration [82, 121] and reporting guidelines [53]. Twenty-five investigated journal policies or author instructions about data sharing [50, 52, 54, 66, 69, 77, 81 with 2 studies, 80, 83, 93 with 2 studies, 96, 101, 103 with 2 studies, 104, 105 with 2 studies, 108, 109, 112, 116, 126, 128]; seven about reporting guidelines [48, 60, 76, 86, 92, 132, 133], and four about preregistration of studies, including trial registration [82, 98, 120, 58]. Two studies from the same author investigated the effect of generic methodological journal guidelines with five years in between the two studies [73,74].

Policy guidelines by funders or government agencies (n=3): Two studies investigated policy issued by funders, more precisely the requirement of the National Institute of Health (NIH) to provide a data sharing plan for large funding grants [102,108]. One study investigated the effect of the US Government's trial registration policy [97].

Reporting guidelines (n=15): Fifteen studies evaluated the effect of reporting guidelines in general [48,59,60,62,63,70,72,88,94,100,104,107,115,130,131]. The baseline for these studies was the publication of a certain reporting checklist or guideline; and the design was either a pre-post design or a post-intervention only design.

Rewards and incentives (n=9): Nine studies investigated the effect of badges for open science or data sharing [65,74,74,83,83,90,111,113]. Two of those investigated the effect of receiving a badge by comparing the amount of data shared from authors who have and have not received a badge, while the other seven compared outcomes between journals issuing badges or not.

Specific tools (n=7): Seven studies looked at specific tools, such as web-based writing or reporting aids (n=3) [55,87,125], workflows and standardization tools (n=2) [95,122], and statistical and automated code cleaning tools (n=2) [61,128].

Other interventions (n=14): Other interventions included, for example, the building of an Open Science training community (n=1) [110], p-value adjustments (n=2) [78,79] and interventions aimed at peer-reviewers (n=2) [64,124]. The complete list can be found in Table 2.

In 61 studies, the party issuing the intervention was the journal, and in 15 studies, academic staff applied the intervention. Academic staff were the population onto whom the intervention was applied in 95 studies, in most cases they were only holistically referred to as “authors”. Two studies were aimed at peer reviewers and one focused specifically on PhD students.

***Interventions not retrieved***

Of the 69 interventions included in our data-extraction sheet, 52 were not investigated in any of the included studies (Supplementary File 4). These can be considered gaps in the evidence we recovered. Examples are interventions such as retractions and funding statements. Also, typical open science interventions, such as data availability, FAIR data, open science ethics, open peer review and open evaluations are lacking in our dataset. We did not retrieve studies that investigated their effect on reproducibility.

**Table 2. List of interventions implemented in the included studies (n = 104).**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Intervention class and subclasses** | **n** | **%** | **Intervention specifics** | **n** | **%** |
| **Open methodology** |  | | | | |
| Open protocol | 1 | 0.96 | publication of protocol | 1 | 0.96 |
| Pre-registration, study registration, analysis plan | 5 | 4.81 | preregistration | 5 | 4.81 |
| Registered reports | 2 | 1.92 | registered reports | 2 | 1.92 |
| Project workflow | 2 | 1.92 | standardization of processes | 2 | 1.92 |
| **Open science policies and guidelines** |  | | | | |
| Rewards and incentives | 9 | 8.65 | badges for data and code sharing | 2 | 1.92 |
| badges for open data | 3 | 2.88 |
| badges for open data and materials | 1 | 0.96 |
| badges for open science | 3 | 2.88 |
| Reporting guidelines, reporting standards | 15 | 14.42 | use of reporting checklist | 3 | 2.88 |
| publication of reporting checklist | 11 | 10.58 |
| publication of methodological guidelines | 1 | 0.96 |
| Policy guidelines  (e.g. of funders/ publishers) | 49 | 47.12 |  | | |
| * on data/code sharing or open science practices |  |  | journal submission guidelines on data/code sharing | 3 | 2.88 |
| journal policy on data/code sharing | 20 | 19.23 |
| ICMJE data sharing statement | 5 | 4.81 |
| journal and funder policies on data/code sharing | 1 | 0.96 |
| funder data sharing policy | 1 | 0.96 |
| journal policy strictness (data/code sharing) | 1 | 0.96 |
| * on research quality |  |  | journal submission guidelines on methodology | 2 | 1.92 |
| * on acceptance and registration before study has started |  |  | journal accepts regardless outcome | 1 | 0.96 |
| journal policy on trial registration | 3 | 2.88 |
| ICMJE proposed trial registration | 2 | 1.92 |
| journal submission guidelines on preregistration | 1 | 0.96 |
| FDA trial registration policy | 1 | 0.96 |
| * on reporting checklists or guidance |  |  | journal endorses reporting checklist | 4 | 3.85 |
| ICMJE endorses reporting guideline | 1 | 0.96 |
| journal requires reporting checklist | 2 | 1.92 |
| journal submission guidelines on reporting guidelines | 1 | 0.96 |
| **Open data and materials** |  | | | | |
| Data availability | 1 | 0.96 | authors shared data | 1 | 0.96 |
| Affirmative sharing declarations | 1 | 0.96 | data sharing statements | 1 | 0.96 |
| **Open educational resources** |  | | | | |
| Training | 2 | 1.92 | training | 2 | 1.92 |
| Other | 1 | 0.96 | results-free review | 1 | 0.96 |
| **Open science tools** |  | | | | |
| Open workflow tools, workflow management systems | 1 | 0.96 | tools | 1 | 0.96 |
| Dockerization, docker | 1 | 0.96 | statistical tool | 1 | 0.96 |
| Other | 5 | 4.81 | automatic code cleaning | 1 | 0.96 |
| standardization of processes | 1 | 0.96 |
| writing tool | 3 | 2.88 |
| **Other interventions** |  | | | | |
| Type1/Type 2 error reduction | 3 | 2.88 | effect-size filter | 1 | 0.96 |
| p-value calibration | 2 | 1.92 |
| Collaborative research | 1 | 0.96 | building a (training) community | 1 | 0.96 |
| Other | 5 | 4.81 | editorial checks/peer review | 3 | 2.88 |
| role models | 1 | 0.96 |
| statistical approach | 1 | 0.96 |

**Outcomes**

Fifteen studies investigated the effect of an intervention of reproducibility or replicability as direct outcomes. One study investigated the effect of data sharing policy on inferential reproducibility [56]. Eight studies investigated whether interventions improved the reproducibility of scientific results: one study investigated the effect of a statistical tool on whether codes could be reproduced [61]; two studies investigated the effect of badges on whether the results of a study could be reproduced without discrepancies [65,83]; two studies investigated the effect of journal policies on whether the results of a study could be reproduced without discrepancies [93,103]; one study investigated the effect of registered reports on whether the same main results with minimal mistakes could be reproduced [106]; one study investigated the effect of workflow tools on whether the same results could be reproduced [95]; and one study investigated the effect of building a training community on whether articles from this community could be reproduced [110]. Two studies investigated the effect of standardization of processes and statistical analyses on replicability of results from different laboratories [51,118]. One article included three studies all about whether statistical inconsistencies could be solved by sharing the data [105], and one study investigated the effect of statistical intervention on the reduction of type-I-errors [89].

Eighty-nine studies assessed the effect of an intervention on a proxy outcome for reproducibility. Thirty-five studies assessed the effect of one or more interventions on transparency of research or methods [48,49,53,55,57,59,60,62–64,70,72–74,76,86–88,92,94,99,100,104,107,115,120,124,125,127,130–132]. This outcome was most commonly assessed via a reporting checklist, such as CONSORT or ARRIVE. One study asked researchers to rate the perceived transparency of their research.

Twenty-six studies assessed the effect of one or more interventions on data-sharing. Of these, ten studies checked whether the data were actually freely available in a repository without first asking for it [50, 56, 73, 74, 77, 80, 96, 108 with 2 studies, 109], four studies only checked whether a statement was made about data availability [69, 84, 102, 119], and twelve studies checked whether data were actually shared after request [67, 81, 85, 101, 103, 104, 112 with 3 studies, 114, 116, 126].

Other proxy outcomes were preregistration and trial data sharing [82,98,121,129]; how transparent researchers rated themselves [68]; and code sharing [54,66,81,93,113]. Some of these outcomes, such as preregistration, have also been included in other studies as interventions.

***Outcomes not retrieved***

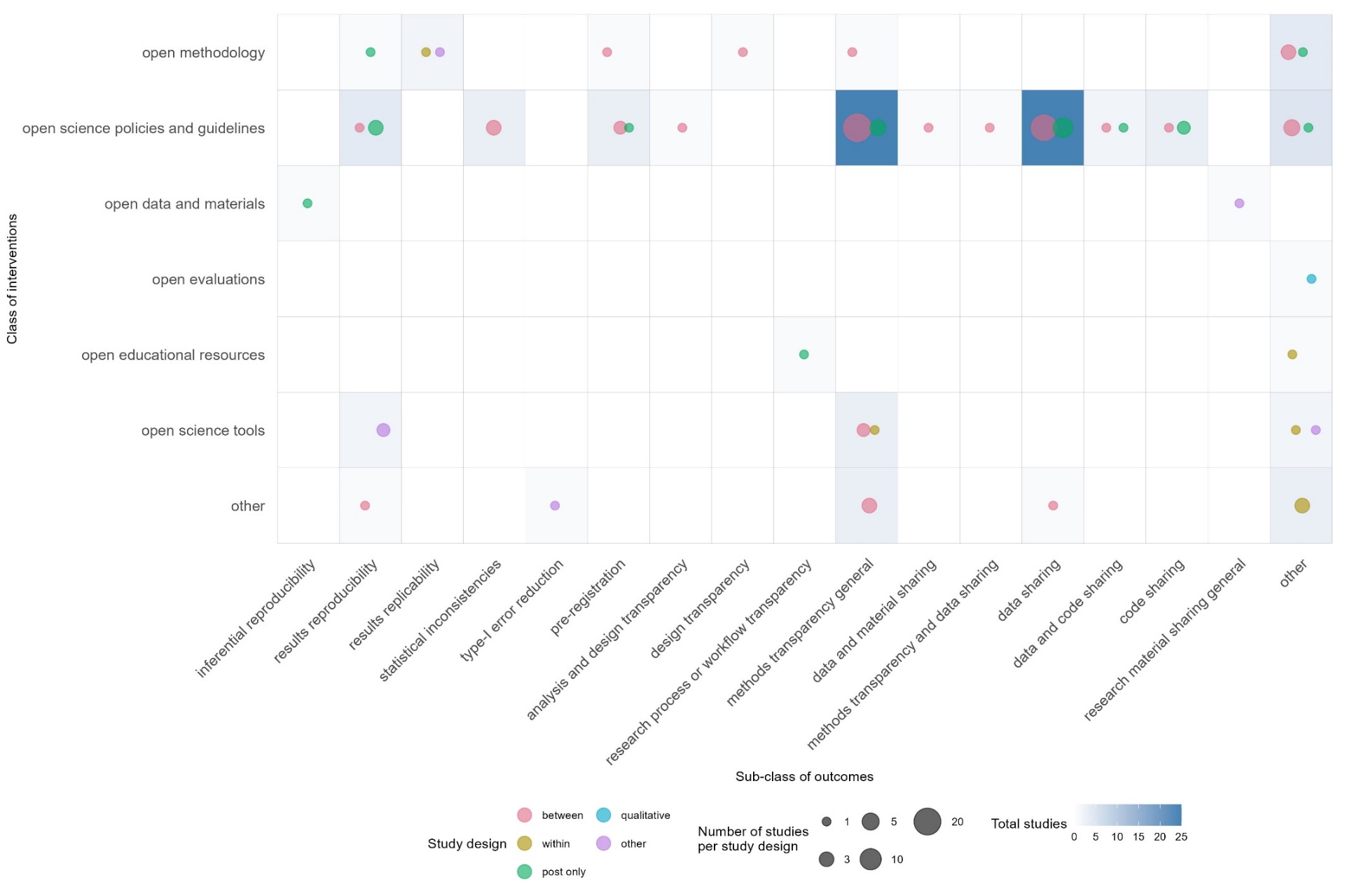
Contrary to our initial list of interventions, most of the items on our initial list of outcomes were investigated. We did not retrieve any studies that directly investigated the effect of interventions on methodological or inferential reproducibility, and we did not retrieve any studies on the proxy outcomes registered reports, post-study peer review, reproducibility checks, data management plans and software sharing.

**Table 3. List of outcomes investigated in the included studies (n=104).**

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome domain** | **Outcomes** | **n** | **%** |
| Reproducibility | Inferential reproducibility | 1 | 0.96 |
| Results reproducibility | 8 | 7.69 |
| Replicability | Results replicability | 2 | 1.92 |
| Other direct outcome | Statistical inconsistencies | 3 | 2.88 |
| Type-I error reduction | 1 | 0.96 |
| Registration status | Pre-registration | 4 | 3.85 |
| Methods transparency | Analysis transparency | 1 | 0.96 |
| Design transparency | 1 | 0.96 |
| Research process or workflow transparency | 1 | 0.96 |
| Methods transparency general | 32 | 30.77 |
| Research material sharing | Material sharing | 1 | 0.96 |
| Data sharing | 29 | 27.88 |
| Code sharing | 3 | 2.88 |
| Research material sharing general | 1 | 0.96 |
| Other proxy outcome | Other | 16 | 15.38 |

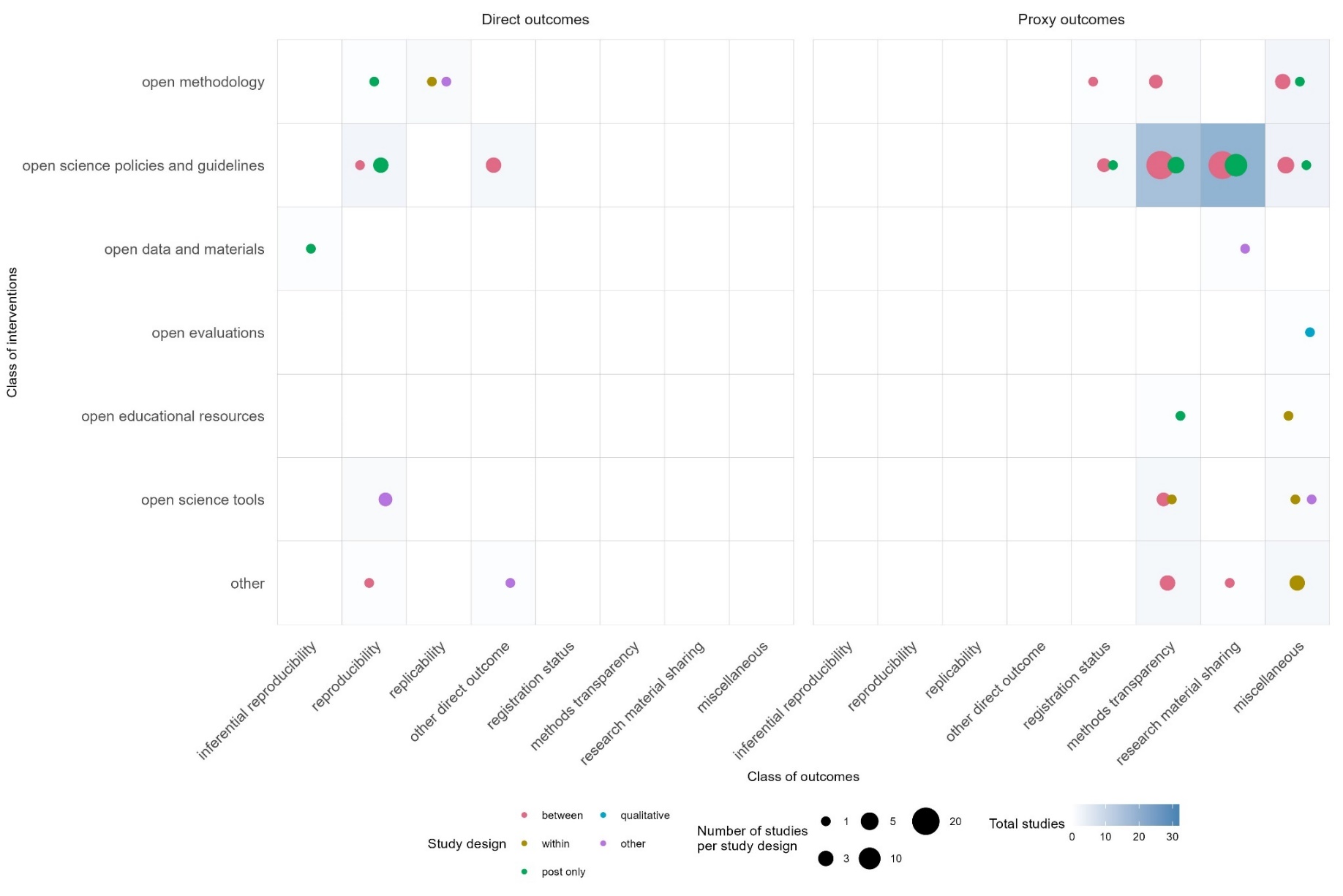
**Evidence gap maps**

The evidence-gap maps show that most research done according to our scoping review is on the effect of open science policies and guidelines for research transparency – measured as the completeness of reporting; and the effect of open science policies and guidelines on data-sharing (Figure 2, Figure 3). None of the studies investigated whether data and materials sharing actually leads to more reproducible research. A few studies did investigate whether shared data could be used to reproduce research findings. Our evidence gap maps show clear gaps. The most notable gap is the lack of studies investigating the effect of open data and materials (data-sharing) on the actual reproducibility and replicability of research. Also, few studies investigated the effect of educational resources and of open evaluations, such as open peer review.

**Figure 2. Evidence gap map of interventions and specific outcomes investigated, with information on study designs.** 

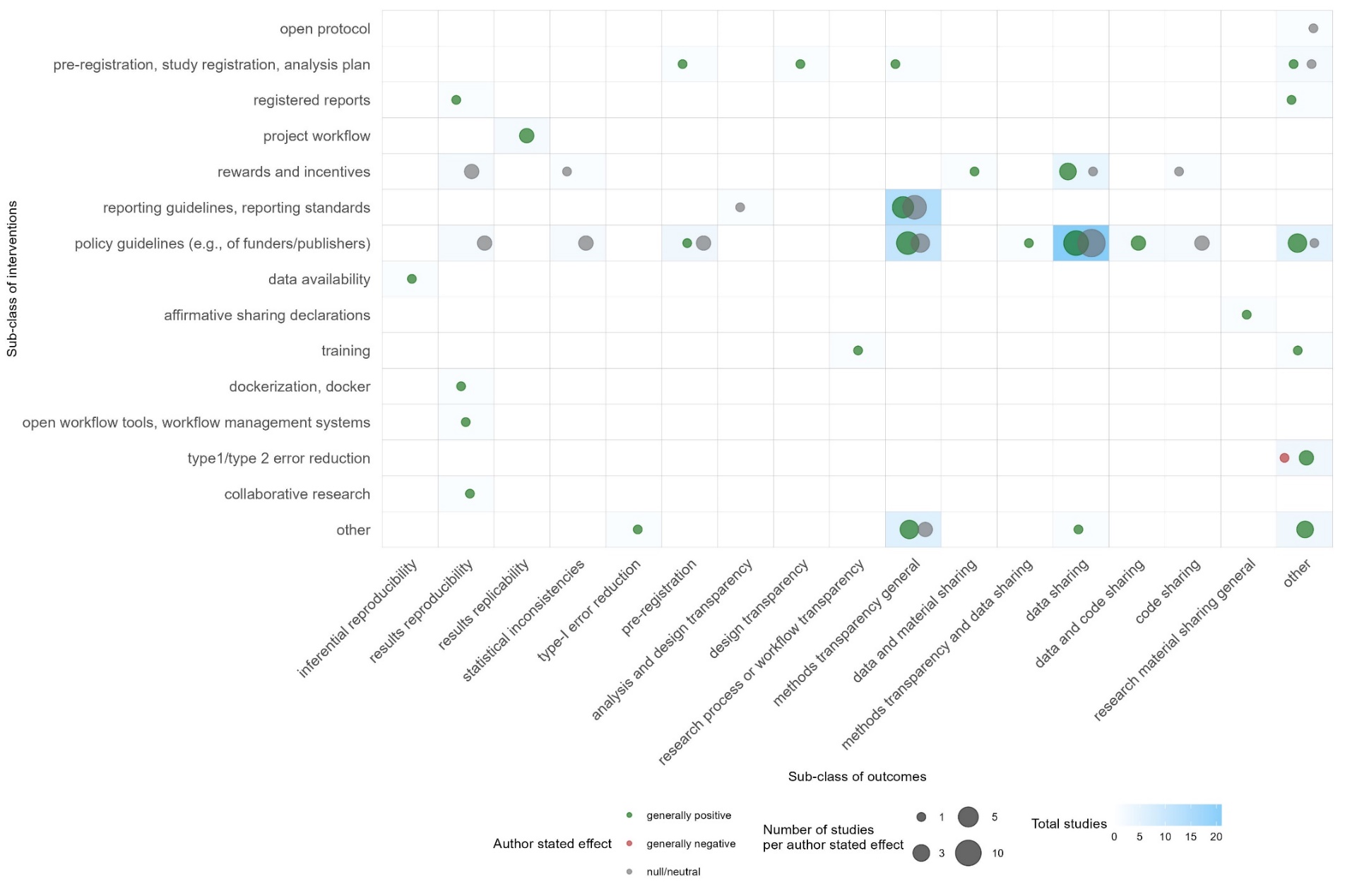
**CAPTION:** Study designs: between = comparative (between-subject comparison) within = comparative (within-subject comparison/repeated measures design) post only = post-intervention (only a post measurement after the implementation of an intervention and the intervention is explicitly mentioned) other = other designs.

**Figure 3. Evidence gap map of interventions and outcome domains investigated, with information on study designs.**



**CAPTION:** The left pane shows direct reproducibility outcomes, while the right pane shows proxy outcomes. Study designs: between = comparative (between-subject comparison) within = comparative (within-subject comparison/repeated measures design) post only = post-intervention (only a post measurement after the implementation of an intervention and the intervention is explicitly mentioned) other = other designs.

**Figure 4. Evidence gap map of interventions and specific outcomes investigated, with information on author-stated effects.**

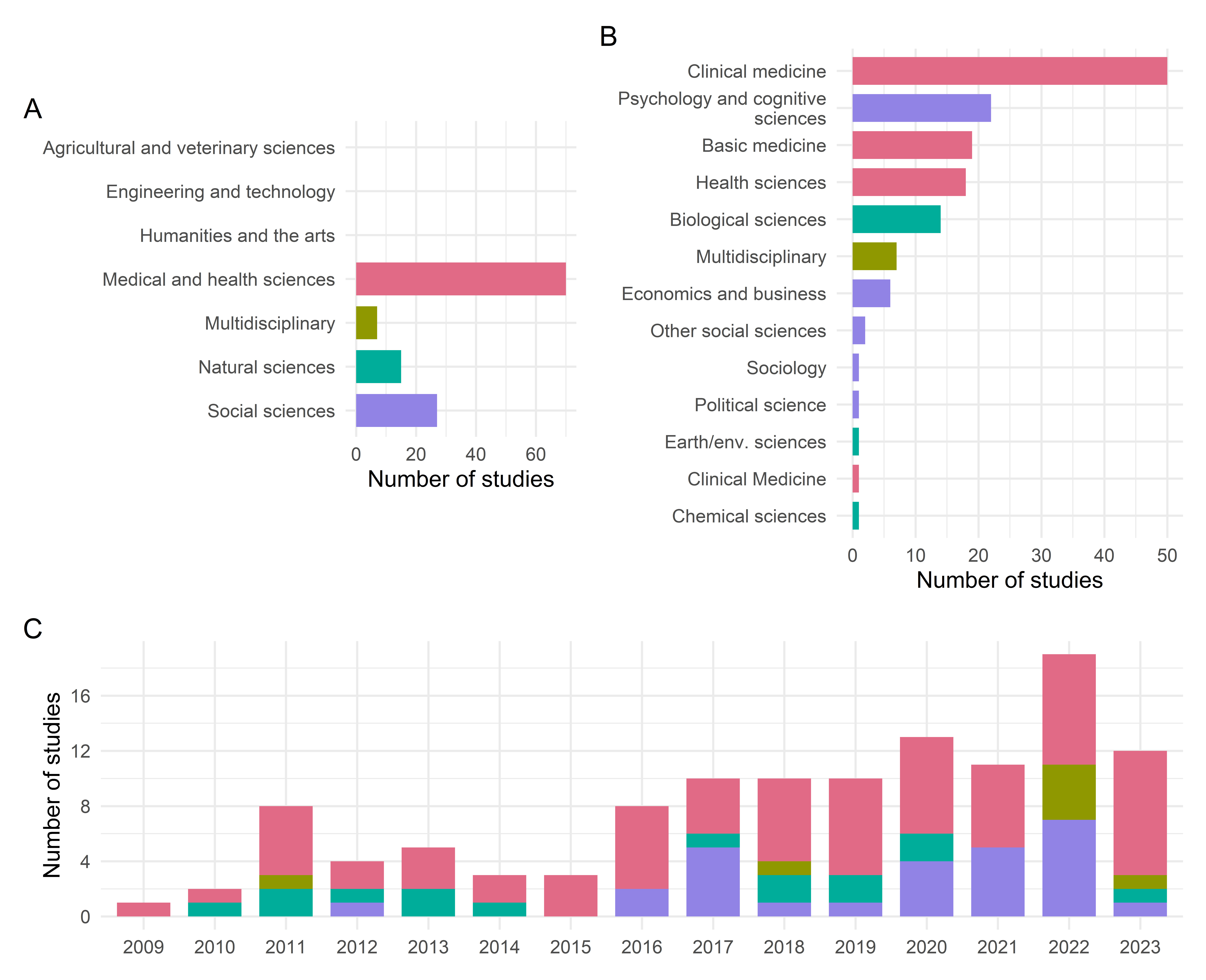


**Author-stated direction of effects**

Of 60 out of 104 research conclusions the authors of the publications rated the effect of the interventions they investigated positively. The authors concluded that these interventions did lead to more reproducibility (or to better methods transparency, more data/code sharing or more preregistration as proxies for reproducibility). Forty-three conclusions were neutral, meaning there was no effect detected or authors concluded that the intervention was not effective. Only one study had a negative conclusion, meaning that the intervention had unintended negative consequences, however only regarding a proxy outcome (p-value calibration leading to an increased type-II error rate) [79].

**Disciplinary scope of interventions**

Figure 5 presents the disciplinary distribution and temporal trends of our included studies. The majority of studies investigating interventions in the medical and health sciences, with significant contributions from the natural sciences and social sciences. Clinical medicine is by far the area most covered, followed by, Psychology/cognitive sciences, Health sciences, Basic medicine, and Biological sciences. Multidisciplinary studies (where three or more disciplines are addressed) as well as Economics/business are also well represented. The temporal distribution of these studies, with the first published in 2009, and running until 2023 (the time of our database search and retrieval), shows a marked increase in the number of studies over time, particularly peaking in 2022 and 2023 (a partial year given the cut-off represented by the date of our database results retrieval in August 2023). The overall trend indicates a growing number (with a notable recent acceleration) of intervention studies to increase reproducibility and replicability across various disciplines, and an overall predominant focus on Clinical medicine, Psychology and cognitive sciences, and other health or biological fields.



**CAPTION:** Disciplinary distribution and temporal trends of included intervention studies, according to the Frascati manual discipline “Fields of Science and Technology“ [134]. **(A)** Number of studies according to Frascati Field of Science and Technology Classification schema top-level ‘Disciplines’. **(B)** Number of studies according to Frascati second-level ‘Knowledge Fields’, with respective top-level categories indicated by colour. **(C)** Number of studies over time, with respective top-level categories indicated by colour. Note that each study may cover more than one Discipline/Field of Knowledge. In these cases, the studies are counted fully for each respective Discipline/Field of Knowledge, and hence overall numbers add up to more than our included number of studies.

**Discussion**

The aim of this scoping review was to gain an overview over formal investigations of interventions to enhance reproducibility and replicability in research, with an emphasis on those related to Open Science. Of the 104 distinct studies we included, only 15 directly measured the effect of an intervention on reproducibility or replicability, while the other research questions addressed a proxy outcome, such as data sharing, methods transparency or preregistration. Thirty research questions were non-comparative and 27 were comparative but cross-sectional, precluding any causal inference. Despite studies investigating a range of interventions, employing a range of intervention methods, and addressing various outcomes, nonetheless our findings indicate that in general the evidence-base of which Open Science interventions improve reproducibility of research remains remarkably limited in many respects.

## ***State of the evidence on effectiveness of interventions***

Despite an ongoing conversation throughout academia regarding the importance of reproducibility and replicability, and the need to intervene to enable more robust and transparent research practices, studies testing interventions in these areas remain relatively scarce. Save for two larger clusters of studies focused on efficacy of data sharing policies and reporting guidelines, many prominently discussed areas of intervention (e.g., training, pre-registration, badges) have been investigated by fewer than 10 studies, often with sub-optimal designs. Regarding outcomes, only 15 studies focused on effects related to reproducibility/replicability directly, with others targeting proxy outcomes mostly related to transparency of reporting or sharing of research materials. While the latter are necessary building-blocks for enabling reproducibility, their study as outcomes does not directly measure any effects upon the robustness of results. Interventions that did focus on reproducibility/replication directly had mixed results, highlighting the need for experimental research specifically designed to test interventions under controlled conditions to provide more definitive evidence. However, while observational and non-randomised designs were common, we identified only a limited number of experimental studies, including six randomized controlled trials.

As many proposed interventions in this space may increase workload and costs for researchers, editors and other stakeholders, it is important that we know these resources are being spent on practices that are evidenced as being effective. We assessed whether the authors of the included articles concluded that they were effective or not. Negative consequences of interventions were basically not reported (e.g., we did not find that something led to less data sharing compared to usual). As this is a scoping review aimed at a general overview of the field and encompasses a wide variety of cross-disciplinary research designs, we did not formally assess the methodological quality of the included research and therefore are not able to firmly draw conclusions about the quality of evidence for the interventions retrieved. However, most included studies used a cross-sectional design or lacked a control group and are by design unsuitable to assess the effect of an intervention.

Lack of research under controlled conditions and a focus on proxy measures rather than reproducibility/replicability themselves may impact the overall robustness of evidence in this area. This limits the extent to which we can currently conclusively discern whether interventions aimed at improving research reproducibility and replicability are actually working. Still, most authors (60 out of 104) concluded positively about the effect of the intervention tested. The overwhelmingly positive to neutral nature of the author-stated claims regarding efficacy found in our review may be ironically indicative that meta-research at present may suffer from possible selective reporting and publication bias [139,140,141] or overoptimistic interpretation of the results [135]. However, it may also be an indication of the causal links between some interventions and reproducibility, such as that using a reporting checklist is likely to lead to (sometimes slightly) better reporting.

## ***Disciplinary coverage and methodological challenges***

Alarm about perceptions of a “reproducibility crisis” were initially fuelled by concerns from a few epistemic communities, especially the health and behavioural sciences. Our findings indicate that research into reproducibility has focused research within these communities. This may indicate a need to explore further evidence for interventions to impact reproducibility in other areas. It could also suggest that other fields do not face the same degree of challenges to reproducibility, are less engaged with these issues, or find reproducibility to be less relevant for their research (for example, in the areas of humanities or qualitative research [25,34,136]).

The limited evidence, with gaps in some areas and complete lack of findings in others, could be attributed to several different causes. Although the debate is becoming increasingly prominent, research that is conducted in this area possibly lags behind what is discussed. Considerable time might pass before explicit calls for research can be met [44] and funding is made available to allow research to meet the scale and scope required of these proposals. Also, poorly evidenced claims on the effectiveness of some intervention, tool or practice might often be repeated, but then not actually tested empirically.

***Limitations***

The primary method for searching articles involved using databases with strict selection criteria. Despite supplementing this with snowballing and searches of secondary literature, it remains quite possible that our review did not include all the available evidence. Of note here, is that although we included all languages, the language-of-publication of the overwhelming majority of our included articles is English. In addition, although our study team encompassed many disciplines, nonetheless our own understandings of reproducibility undoubtedly shaped the search strategy and data-extraction.

Discussions about terms and definitions accompanied us throughout the process due to the immense variety in language around reproducibility and related concepts. This was due to the multidisciplinary nature of our team and the broad scope of our review. On top of that, the reporting of some of the included studies left room for multiple interpretations. Barriers like jargon and unclear methodological specification may have limited our understanding of some findings. Furthermore, we carry our own beliefs and expected to find a scarcity of evidence for most interventions investigated. Hence, despite our best efforts to maintain balance, the combination of our specific perspectives and the width of the included studies have likely shaped our study findings.

Another major limitation is that the data-extraction was done by one person for each included study, especially in combination with the lack of clarity of the included studies and the multidisciplinary team. Therefore, any doubts about what we read were discussed in our weekly meetings, and the last author checked all the extracted data before the analyses were done.

***Future directions and implications for policy and practice***

Greater understanding of the meanings of reproducibility and replicability across contexts may facilitate the understanding of what works in improving levels of reproducibility/replicability and in which circumstances. Therefore, the next step should be to build on this work to investigate and compare in more detail the studies we have identified to synthesise the knowledge contained therein at a more granular level. Questions include: Which specific interventions show most promise as being effective and therefore merit further investigation? Is there a difference in effect when policies are mandatory versus optional? What characterises the interventions that are reported to potentially be effective? How does the efficacy of different interventions vary across different stakeholders or implementing institutions? As these questions require an in-depth qualitative analysis that would go beyond the aim of this scoping review, we will address some of these questions in a follow-up study.

Our findings, perhaps above all else, demonstrate a substantial need within the meta-research community for more thorough and rigorous investigations into the effectiveness of interventions aiming to increase levels of reproducibility and replicability. It is striking that researchers, policymakers, journals and others are often content to implement large-scale changes without rigorous exploration of their efficacy. More specifically, we call for increased use of standardised and robust experimental designs. Implementing such studies will often involve nuanced approaches involving various stakeholders—including institutions, journals, funders, and researchers—who traditionally may not be accustomed to scrutinizing their processes through rigorous experimentation. These groups must become more receptive to subjecting their processes to investigation, as well as share data for experimental or observational research [137,138]. Given the previously mentioned observation about possibly overoptimistic reporting or possible selective reporting, we highlight the need for more “red-teaming” [142] of interventions – with researchers less invested in positive outcomes participating.

Given the crucial need to better understand how “epistemic diversity” influences the meaning, relevance and feasibility of reproducibility research [25,27], a corollary of our findings is that further exploring the effectiveness of reproducibility interventions across diverse research contexts is essential. This could be facilitated by global or multi-stakeholder collaborations aimed at developing a common framework for assessing reproducibility and replicability, thereby enhancing the generalisability of findings where possible. Furthermore, expanding the range of disciplines involved in reproducibility research could provide broader insights and innovations, especially regarding which interventions can be implemented at a cross-disciplinary level (leveraging economies of scale) and which require detailed discipline-specific applications.

All this requires funding and resources, of course. Policy-makers, funders and institutions, but also publishers and scholarly societies, should consider supporting these lines of inquiry through dedicated funding streams and in-kind access to data and systems for experimentation. Such support would foster more interdisciplinary approaches and help break down the silos that often separate research communities. Towards the latter, further recognising and supporting 'research on research' as a distinct field with its own conferences, journals and funding lines will also help. Doing so may enable immediate returns on investments through reducing wasted resources dedicated to ineffective interventions.

**Ethics and consent**

Ethical approval and consent were not required.

**Data availability**

All materials, including the data extraction sheets, data and code are available on OSF (DOI [[10.17605/OSF.IO/7EF5H](http://dx.doi.org/10.17605/OSF.IO/7EF5H))](http://dx.doi.org/10.17605/OSF.IO/7EF5H)

**Funding statement**

This work was supported by the projects OSIRIS and TIER2, funded by the European Union’s Horizon Europe research and innovation program under grant agreement No. 101094725 (OSIRIS) and 101094817 (TIER2). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (ERA). Neither the European Union nor the ERA can be held responsible for them.

**CRediT Author Statement**

Leonie Dudda (ORCID 0009-0000-0759-7354): Methodology, investigation, writing (original draft), project administration.

Eva Kormann (ORCID 0009-0005-5680-365): Investigation, writing (review and editing)

Magdalena Kozula (ORCID 0000-0003-0333-4279): Investigation, writing (review and editing)

Nicholas J. DeVito (ORCID 0000-0001-8286-1995): Conceptualization, investigation, writing (review and editing)

Thomas Klebel (ORCID 0000-0002-7331-4751): Software, investigation, writing (review and editing)

Ayu P.M. Dewi: Software, formal analysis, investigation, writing (review and editing), visualization

René Spijker (ORCID 0000-0003-4445-1201): Investigation, writing (review and editing), project organization.

Inge Stegeman (ORCID 0000-0001-5154-7178): Conceptualization, methodology, writing (review and editing), supervision, funding acquisition.

Veerle Van den Eynden (ORCID 0000-0003-2542-2747): Conceptualization, investigation, data curation, writing (original draft, review and editing), supervision, resources, funding acquisition.

Tony Ross-Hellauer (ORCID 0000-0003-4470-7027): Conceptualization, investigation, writing (original draft, review and editing), supervision, resources, funding acquisition.

Mariska M.G. Leeflang (ORCID 0000-0001-5960-0471): Conceptualization, methodology, formal analysis, investigation, data curation, writing (original draft, review and editing), supervision, funding acquisition.

**Acknowledgements**

We gladly acknowledge the contributions of members of OSIRIS and TIER2 in comments on the draft protocol (Patrick Onghena, Florian Naudet, Rita Banzi, Maddalena Fratelli, Monika Varga, Yuri Gelsleichter), literature screening (Gowri Gopalakrisha, Stefania Amodeo, Simone Kopeinik), and helpful comments on an advanced draft of this manuscript (Florian Naudet, Hynek Roubik, David Moher, Olavo Boher Amaral, Massimo Grassi).

**References**

1. Baker M. 1,500 scientists lift the lid on reproducibility. Nature. 2016;533: 452–454. doi:10.1038/533452a

2. Vazire S. Implications of the Credibility Revolution for Productivity, Creativity, and Progress. Perspect Psychol Sci. 2018;13: 411–417. doi:10.1177/1745691617751884

3. Ioannidis JPA. Why Most Published Research Findings Are False. PLOS Med. 2005;2: e124. doi:10.1371/journal.pmed.0020124

4. Nosek BA, Hardwicke TE, Moshontz H, Allard A, Corker KS, Dreber A, et al. Replicability, Robustness, and Reproducibility in Psychological Science. Annu Rev Psychol. 2022;73: 719–748. doi:10.1146/annurev-psych-020821-114157

5. Bishop D. Rein in the four horsemen of irreproducibility. Nature. 2019;568: 435–435. doi:10.1038/d41586-019-01307-2

6. Phillips PWB, Castle D, Smyth SJ. Evidence-based policy making: determining what is evidence. Heliyon. 2020;6: e04519. doi:10.1016/j.heliyon.2020.e04519

7. Saltelli A, Giampietro M. What is wrong with evidence based policy, and how can it be improved? Futures. 2017;91: 62–71. doi:10.1016/j.futures.2016.11.012

8. Wingen T, Berkessel JB, Englich B. No Replication, No Trust? How Low Replicability Influences Trust in Psychology. Soc Psychol Personal Sci. 2020;11: 454–463. doi:10.1177/1948550619877412

9. National Academies of Sciences E, Division H and M, Services B on HC, Policy B on HS, Health R on G and P, Forum NCP, et al. Transparency and Trust. Enhancing Scientific Reproducibility in Biomedical Research Through Transparent Reporting: Proceedings of a Workshop. National Academies Press (US); 2020. Available: https://www.ncbi.nlm.nih.gov/books/NBK556622/

10. Goodman SN, Fanelli D, Ioannidis JPA. What does research reproducibility mean? Sci Transl Med. 2016;8: 341ps12-341ps12. doi:10.1126/scitranslmed.aaf5027

11. Gelman A. Essay: The Experiments Are Fascinating. But Nobody Can Repeat Them. The New York Times. 20 Nov 2018. Available: https://www.nytimes.com/2018/11/19/science/science-research-fraud-reproducibility.html. Accessed 21 May 2024.

12. Feilden T. Most scientists “can’t replicate studies by their peers.” BBC News. 22 Feb 2017. Available: https://www.bbc.com/news/science-environment-39054778. Accessed 21 May 2024.

13. Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015;349: aac4716. doi:10.1126/science.aac4716

14. Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov. 2011;10: 712–712. doi:10.1038/nrd3439-c1

15. Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, et al. Investigating the replicability of preclinical cancer biology. Elife. 2021;10: null. doi:10.7554/eLife.71601

16. Camerer CF, Dreber A, Forsell E, Ho T-H, Huber J, Johannesson M, et al. Evaluating replicability of laboratory experiments in economics. Science. 2016;351: 1433–1436. doi:10.1126/science.aaf0918

17. Camerer CF, Dreber A, Holzmeister F, Ho TH, Huber J, Johannesson M, et al. Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nat Hum Behav. 20180827th ed. 2018;2: 637–644. doi:10.1038/s41562-018-0399-z

18. Six factors affecting reproducibility in life science research and how to handle them. [cited 7 May 2024]. Available: https://www.nature.com/articles/d42473-019-00004-y

19. Munafo MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, du Sert NP, et al. A manifesto for reproducible science. Nat Hum Behav. 2017;1: 0021. doi:10.1038/s41562-016-0021

20. National Academies of Sciences E, Affairs P and G, Committee on Science E, Information B on RD and, Sciences D on E and P, Statistics C on A and T, et al. Understanding Reproducibility and Replicability. Reproducibility and Replicability in Science. National Academies Press (US); 2019. Available: https://www.ncbi.nlm.nih.gov/books/NBK547546/

21. Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JPA. Public Availability of Published Research Data in High-Impact Journals. PLoS ONE. 2011;6: e24357. doi:10.1371/journal.pone.0024357

22. Stodden V. Assessing Reproducibility : An Astrophysical Example of Computational Uncertainty in the HPC Context. 2018. Available: https://www.semanticscholar.org/paper/c2321c16a12b0a71ab99214a054a60bd2655a505

23. Athena RC, Directorate-General for Research and Innovation (European Commission), Know-Center, PPMI. Assessing the reproducibility of research results in EU Framework Programmes for Research: final report. Publications Office of the European Union; 2022. Available: https://data.europa.eu/doi/10.2777/186782

24. Friesike S, Schildhauer T. Open Science: Many Good Resolutions, Very Few Incentives, Yet. In: Welpe IM, Wollersheim J, Ringelhan S, Osterloh M, editors. Incentives and Performance. Cham: Springer International Publishing; 2015. pp. 277–289. doi:10.1007/978-3-319-09785-5\_17

25. Leonelli S. Rethinking Reproducibility as a Criterion for Research Quality. Including a Symposium on Mary Morgan: Curiosity, Imagination, and Surprise. Emerald Publishing Limited; 2018. pp. 129–146. doi:10.1108/S0743-41542018000036B009

26. Leonelli S. Open Science and Epistemic Diversity: Friends or Foes? Philos Sci. 2022;89: 991–1001. doi:10.1017/psa.2022.45

27. Ulpts S, Schneider JW. Knowledge Production Modes: The Relevance and Feasibility of ‘Reproducibility.’ OSF; 2023. doi:10.31222/osf.io/ujnd9

28. Penders B, Holbrook JB, de Rijcke S. Rinse and Repeat: Understanding the Value of Replication across Different Ways of Knowing. Publications. 2019;7: 52. doi:10.3390/publications7030052

29. Peels R. Replicability and replication in the humanities. Res Integr Peer Rev. 2019;4: 2. doi:10.1186/s41073-018-0060-4

30. Peels R, Bouter L. Humanities need a replication drive too. Nature. 2018;558: 372–372. doi:10.1038/d41586-018-05454-w

31. Peels R, Bouter L. The possibility and desirability of replication in the humanities. Palgrave Commun. 2018;4: 1–4. doi:10.1057/s41599-018-0149-x

32. Makel MC, Meyer MS, Simonsen MA, Roberts AM, Plucker JA. Replication is relevant to qualitative research. Educ Res Eval. 2022 [cited 7 May 2024]. Available: https://www.tandfonline.com/doi/abs/10.1080/13803611.2021.2022310

33. de Rijcke S, Penders B. Resist calls for replicability in the humanities. Nature. 2018;560: 29–29. doi:10.1038/d41586-018-05845-z

34. Tuval-Mashiach R. Is replication relevant for qualitative research? Qual Psychol. 2021;8: 365–377. doi:10.1037/qup0000217

35. Pownall M. Is replication possible in qualitative research? A response to Makel et al. (2022). Educ Res Eval. 2024;29: 104–110. doi:10.1080/13803611.2024.2314526

36. The Registration of Observational Studies—When Metaphors Go Bad. Epidemiology. 2010;21: 607–609. doi:10.1097/EDE.0b013e3181eafbcf

37. Dal-Ré R, Ioannidis JP, Bracken MB, Buffler PA, Chan A-W, Franco EL, et al. Making Prospective Registration of Observational Research a Reality. Sci Transl Med. 2014;6: 224cm1-224cm1. doi:10.1126/scitranslmed.3007513

38. Hansford H. The limited role of preregistration in observational studies. In: Harrison Hansford [Internet]. 4 Aug 2022 [cited 7 May 2024]. Available: https://harrisonhansford.home.blog/2022/08/04/the-limited-role-of-preregistration-in-observational-studies/

39. Coffman LC, Niederle M. Pre-analysis Plans Have Limited Upside, Especially Where Replications Are Feasible. J Econ Perspect. 2015;29: 81–98. doi:10.1257/jep.29.3.81

40. Haven T, Van Grootel L. Preregistering qualitative research. Account Res. 2019;26: 229–244. doi:10.1080/08989621.2019.1580147

41. Goodman A, Pepe A, Blocker AW, Borgman CL, Cranmer K, Crosas M, et al. Ten Simple Rules for the Care and Feeding of Scientific Data. Bourne PE, editor. PLoS Comput Biol. 2014;10: e1003542. doi:10.1371/journal.pcbi.1003542

42. Ayris P, Lopez de San Román A, Maes K, Labastida I. Open science and its role in universities: A roadmap for cultural change. Leuven LERU Off Retrieved Dec. 2018;13: 2019.

43. Hernández JA, Colom M. Repeatability, Reproducibility, Replicability, Reusability (4R) in Journals’ Policies and Software/Data Management in Scientific Publications: A Survey, Discussion, and Perspectives. arXiv; 2023. doi:10.48550/arXiv.2312.11028

44. Suls J, Rothman AJ, Davidson KW. Now is the time to assess the effects of open science practiceswith randomized control trials. Am Psychol. 2022;77: 467–475. doi:10.1037/amp0000871

45. Dudda LA, Kozula M, Ross-Hellauer T, Kormann E, Spijker R, DeVito N, et al. Scoping review and evidence mapping of interventions aimed at improving reproducible and replicable science: Protocol. Open Res Eur. 2023;3: 179. doi:10.12688/openreseurope.16567.1

46. Tricco AC, Lillie E, Zarin W, O’Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169: 467–473. doi:10.7326/M18-0850

47. Directorate-General for Research and Innovation (European Commission), Baker L, Cristea IA, Errington TM, Jaśko K, Lusoli W, et al. Reproducibility of scientific results in the EU: scoping report. Publications Office of the European Union; 2020. Available: https://data.europa.eu/doi/10.2777/341654

48. Agha R A, Borrelli M R, Farwana R, Kusu-Orkar T, Millip M C, Thavayogan R, et al. Impact of the PROCESS guideline on the reporting of surgical case series: A before and after study. Int J Surg. 2017;45: 92–97. doi:10.1016/j.ijsu.2017.07.079

49. Agha R A, Fowler A J, Limb C, Whitehurst K, Coe R, Sagoo H, et al. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: A before and after study. Int J Surg. 2016;30: 169–172. doi:10.1016/j.ijsu.2016.04.032

50. Alsheikh-Ali A A, Qureshi W, Al-Mallah M H, Ioannidis J P. Public availability of published research data in high-impact journals. 2011;1: e24357.

51. Arroyo-Araujo M, Graf R, Maco M, van Dam E, Schenker E, Drinkenburg W, et al. Reproducibility via coordinated standardization: a multi-center study in a Shank2 genetic rat model for Autism Spectrum Disorders. Sci Rep. 2019;9: 11602. doi:10.1038/s41598-019-47981-0

52. Askarov Z, Doucouliagos A, Doucouliagos H, Stanley T D. The Significance of Data-Sharing Policy. J Eur Econ Assoc. 2022. doi:10.1093/jeea/jvac053

53. Audet S, Doyle C, Lemieux C, Lemieux J. Examining the adherence to CONSORT guidelines and the reporting of the enrollment process in clinical oncology randomized controlled trials: A review of trials published between 2013 and 2015. J Clin Oncol. 2017;35. doi:10.1200/jco.2017.35.15\_suppl.e14026

54. Auspurg K, Recker J. More openness in research? An evaluation of open science measures at the Journal of Sociology. Z Soziol. 2020;49: 1–9. doi:10.1515/zfsoz-2020-0001

55. Barnes C, Boutron I, Giraudeau B, Porcher R, Altman DG, Ravaud P. Impact of an online writing aid tool for writing a randomized trial report: the COBWEB (Consort-based WEB tool) randomized controlled trial. BMC Med. 2015;13: 221. doi:10.1186/s12916-015-0460-y

56. Bergeat D, Lombard N, Gasmi A, Le Floch B, Naudet F. Data Sharing and Reanalyses Among Randomized Clinical Trials Published in Surgical Journals Before and After Adoption of a Data Availability and Reproducibility Policy. JAMA Netw Open. 2022;5: e2215209. doi:10.1001/jamanetworkopen.2022.15209

57. Blanco David, Schroter Sara, Aldcroft Adrian, Moher David, Boutron Isabelle, Kirkham Jamie J, et al. Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial. BMJ Open. 2020;10: e036799. doi:10.1136/bmjopen-2020-036799

58. Blanco-Perez C, Brodeur A. Publication Bias and Editorial Statement on Negative Findings. Econ J. 2020;130: 1226–1247. doi:10.1093/ej/ueaa011

59. Can O S, Yilmaz A A, Hasdogan M, Alkaya F, Turhan S C, Can M F, et al. Has the quality of abstracts for randomised controlled trials improved since the release of Consolidated Standards of Reporting Trial guideline for abstract reporting? A survey of four high-profile anaesthesia journals. Eur J Anaesthesiol. 2011;28(7): 485–492. doi:10.1097/eja.0b013e32833fb96f

60. Caron J E, March J K, Cohen M B, Schmidt R L. A Survey of the Prevalence and Impact of Reporting Guideline Endorsement in Pathology Journals. Am J Clin Pathol. 2017;148: 314–322. doi:10.1093/ajcp/aqx080

61. Chan C H, Schoch D. rang: Reconstructing reproducible R computational environments. PLoS ONE Electron Resour. 2023;18: e0286761. doi:10.1371/journal.pone.0286761

62. Chow J T. Y, Turkstra T P, Yim E, Jones P M. The degree of adherence to CONSORT reporting guidelines for the abstracts of randomised clinical trials published in anaesthesia journals: A cross-sectional study of reporting adherence in 2010 and 2016. Eur J Anaesthesiol. 2018; 942–948. doi:10.1097/eja.0000000000000880

63. Clayson P E, Carbine K A, Baldwin S A, Larson M J. Methodological reporting behavior, sample sizes, and statistical power in studies of event-related potentials: Barriers to reproducibility and replicability. Psychophysiology. 2019;56: e13437. doi:10.1111/psyp.13437

64. Cobo E, Cortes J, Ribera J M, Cardellach F, Selva-O’Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. 2011;1: d6783.

65. Crüwell S, Apthorp D, Baker B J, Colling L, Elson M, Geiger S J, et al. What’s in a Badge? A Computational Reproducibility Investigation of the Open Data Badge Policy in One Issue of Psychological Science. Psychol Sci. 2023;34: 512–522. doi:10.1177/09567976221140828

66. Culina A, van den Berg I, Evans S, Sanchez-Tojar A. Low availability of code in ecology: A call for urgent action. Plos Biol. 2020;18: e3000763. doi:10.1371/journal.pbio.3000763

67. Danchev V, Min Y, Borghi J, Baiocchi M, Ioannidis J P. A. Evaluation of Data Sharing After Implementation of the International Committee of Medical Journal Editors Data Sharing Statement Requirement. JAMA Netw Open. 2021;4: e2033972. doi:10.1001/jamanetworkopen.2020.33972

68. Deardorff A. Assessing the impact of introductory programming workshops on the computational reproducibility of biomedical workflows. PLoS ONE Electron Resour. 2020;15: e0230697. doi:10.1371/journal.pone.0230697

69. Federer L M, Belter C W, Joubert D J, Livinski A, Lu Y L, Snyders L N, et al. Data sharing in PLOS ONE: An analysis of Data Availability Statements. PLoS ONE Electron Resour. 2018;13: e0194768. doi:10.1371/journal.pone.0194768

70. Fernandes J G, Franco N H, Grierson A J, Hultgren J, Furley A J. W, Olsson I A. S. Methodological standards, quality of reporting and regulatory compliance in animal research on amyotrophic lateral sclerosis: a systematic review. BMJ Open Sci. 2019;3: e000016. doi:10.1136/bmjos-2018-000016

71. Findley M G, Jensen N M, Malesky E J, Pepinsky T B. Can Results-Free Review Reduce Publication Bias? The Results and Implications of a Pilot Study. Comp Polit Stud. 2016;49: 1667–1703. doi:10.1177/0010414016655539

72. Fuller T, Peters J, Pearson M, Anderson R. Impact of the transparent reporting of evaluations with nonrandomized designs reporting guideline: Ten years on. Am J Public Health. 2014;104: e110–e117. doi:10.2105/ajph.2014.302195

73. Giofrè D, Boedker I, Cumming G, Rivella C, Tressoldi P. The influence of journal submission guidelines on authors’ reporting of statistics and use of open research practices: Five years later. Behav Res Methods. 2022;17: 17. doi:10.3758/s13428-022-01993-3

74. Giofrè D, Cumming G, Fresc L, Boedker I, Tressoldi P. The influence of journal submission guidelines on authors’ reporting of statistics and use of open research practices. PLoS ONE Electron Resour. 2017;12: e0175583. doi:10.1371/journal.pone.0175583

75. Gkiouras K, Choleva M E, Verrou A, Goulis D G, Bogdanos D P, Grammatikopoulou M G. A Meta-Epidemiological Study of Positive Results in Clinical Nutrition Research: The Good, the Bad and the Ugly of Statistically Significant Findings. Nutrients. 2022;14: 04. doi:10.3390/nu14235164

76. Goldacre B, Drysdale H, Dale A, Milosevic I, Slade E, Hartley P, et al. COMPare: a prospective cohort study correcting and monitoring 58 misreported trials in real time. Trials Electron Resour. 2019;20: 118. doi:10.1186/s13063-019-3173-2

77. Gorman D M. Availability of Research Data in High-Impact Addiction Journals with Data Sharing Policies. Sci Eng Ethics. 2020;26: 1625–1632. doi:10.1007/s11948-020-00203-7

78. Gregori J, Villarreal L, Sanchez A, Baselga J, Villanueva J. An effect size filter improves the reproducibility in spectral counting-based comparative proteomics. J Proteomics. 2013;95: 55–65. doi:10.1016/j.jprot.2013.05.030

79. Gruber S, Tchetgen Tchetgen, E. Limitations of empirical calibration of p-values using observational data. 2016;1: 3869–3882.

80. Hamilton D G, Hong K, Fraser H, Rowhani-Farid A, Fidler F, Page M J. Prevalence and predictors of data and code sharing in the medical and health sciences: systematic review with meta-analysis of individual participant data. BMJ. 2023;382: e075767. doi:10.1136/bmj-2023-075767

81. Hamilton D G, Page M J, Finch S, Everitt S, Fidler F. How often do cancer researchers make their data and code available and what factors are associated with sharing? BMC Med. 2022;20: 438. doi:10.1186/s12916-022-02644-2

82. Hardt J L. S, Metzendorf M I, Meerpohl J J. Surgical trials and trial registers: A cross-sectional study of randomized controlled trials published in journals requiring trial registration in the author instructions. Trials. 2013;14: 407. doi:10.1186/1745-6215-14-407

83. Hardwicke T E, Bohn M, MacDonald K, Hembacher E, Nuijten M B, Peloquin B N, et al. Analytic reproducibility in articles receiving open data badges at the journal Psychological Science: an observational study. R Soc Open Sci. 2021;8: 201494. doi:10.1098/rsos.201494

84. Hardwicke T E, Mathur M B, MacDonald K, Nilsonne G, Banks G C, Kidwell M C, et al. Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal Cognition. R Soc Open Sci. 2018;5: 180448. doi:10.1098/rsos.180448

85. Haven T L, Abunijela S, Hildebrand N. Biomedical supervisors’ role modeling of open science practices. eLife. 2023;12: 22. doi:10.7554/elife.83484

86. 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 Online. 2012;345(7864) (no pagination): e4178. doi:10.1136/bmj.e4178

87. Hopewell Sally, Boutron Isabelle, Altman Douglas G, Barbour Ginny, Moher David, Montori Victor, et al. Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: results of a randomised controlled trial. BMC Med. 2016;14: 199. doi:10.1186/s12916-016-0736-x

88. Howell V, Schwartz A E, O’Leary J D, McDonnell C. The effect of the SQUIRE (Standards of QUality Improvement Reporting Excellence) guidelines on reporting standards in the quality Improvement literature: A before-and-after study. BMJ Qual Saf. 2015;24(6): 400–406. doi:10.1136/bmjqs-2014-003737

89. Jaljuli I, Kafkafi N, Giladi E, Golani I, Gozes I, Chesler E J, et al. A multi-lab experimental assessment reveals that replicability can be improved by using empirical estimates of genotype-by-lab interaction. Plos Biol. 2023;21: e3002082. doi:10.1371/journal.pbio.3002082

90. Kidwell M C, Lazarevic L B, Baranski E, Hardwicke T E, Piechowski S, Falkenberg L S, et al. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. PLoS Biol. 2016;14(5) (no pagination): e1002456. doi:10.1371/journal.pbio.1002456

91. Krawczyk M, Reuben E. (Un)available upon request: field experiment on researchers’ willingness to share supplementary materials. 2012;1: 175–186.

92. Ladd C, Greenough M C, Reddy A K, Garrett E, Peterson A, Pierce A, et al. An evaluation of reporting guidelines and Clinical Trial Registry requirements among anesthesiology journals. J Evid-Based Med. 2023;16: 116–119. doi:10.1111/jebm.12533

93. Laurinavichyute A, Yadav H, Vasishth S. Share the code, not just the data: A case study of the reproducibility of articles published in the Journal of Memory and Language under the open data policy. J Mem Lang. 2022;125: 104332. doi:10.1016/j.jml.2022.104332

94. Liu H, Gielen Mjcam, Bosmans Jwam, Winkens B, Bouvy N D. Inadequate awareness of adherence to ARRIVE guidelines, regarding reporting quality of hernia models repaired with meshes: a systematic review. Hernia. 2022;26: 389–400. doi:10.1007/s10029-020-02351-y

95. Madduri R, Chard K, D’Arcy M, Jung S C, Rodriguez A, Sulakhe D, et al. Reproducible big data science: A case study in continuous FAIRness. PLoS ONE Electron Resour. 2019;14: e0213013. doi:10.1371/journal.pone.0213013

96. Magee A F, May M R, Moore B R. The dawn of open access to phylogenetic data. PLoS ONE Electron Resour. 2014;9: e110268. doi:10.1371/journal.pone.0110268

97. Marin Dos Santos, D H, Atallah A N. FDAAA legislation is working, but methodological flaws undermine the reliability of clinical trials: a cross-sectional study. 2015;1: e1015. doi:10.7717/peerj.1015

98. McGee R G, Su M, Kelly P J, Higgins G Y, Craig J C, Webster A C. Trial registration and declaration of registration by authors of randomized controlled trials. Transplantation. 2011;92(10): 1094–1100. doi:10.1097/tp.0b013e318232baf2

99. Menon J M. L, Ritskes-Hoitinga M, Pound P, van Oort E. The impact of conducting preclinical systematic reviews on researchers and their research: A mixed method case study. PLoS ONE Electron Resour. 2021;16: e0260619. doi:10.1371/journal.pone.0260619

100. Meredith A J, Simeon-Dubach D, Matzke L A, Cheah S, Watson P H. Biospecimen Data Reporting in the Research Literature. Biopreservation Biobanking. 2019;17: 326–333. doi:10.1089/bio.2018.0143

101. Milia N, Congiu A, Anagnostou P, Montinaro F, Capocasa M, Sanna E, et al. Mine, yours, ours? Sharing data on human genetic variation. 2012;1: e37552.

102. Narang C, Ouvina M, Rees C A, Bourgeois F T. Data Sharing for Pediatric Clinical Trials Funded by the US National Institutes of Health. JAMA Netw Open. 2023;6: e2325342. doi:10.1001/jamanetworkopen.2023.25342

103. Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: Survey of studies published in the BMJ and PLOS Medicine. BMJ Online. 2018;360: k400. doi:10.1136/bmj.k400

104. Nguyen P Y, Kanukula R, McKenzie J E, Alqaidoom Z, Brennan S E, Haddaway N R, et al. Changing patterns in reporting and sharing of review data in systematic reviews with meta-analysis of the effects of interventions: Cross sectional meta-research study. BMJ. 2022; e072428. doi:10.1136/bmj-2022-072428

105. Nuijten M B, Borghuis J, Veldkamp C L. S, Dominguez-Alvarez L, Van Assen M A. L. M, Wicherts J M. Journal data sharing policies and statistical reporting inconsistencies in psychology. Collabra Psychol. 2017;3: 31. doi:10.1525/collabra.102

106. Obels Pepijn, Lakens Daniel, Coles Nicholas A, Gottfried Jaroslav, Green Seth A. Analysis of open data and computational reproducibility in registered reports in psychology. Adv Methods Pract Psychol Sci. 2020;3: 229–237. doi:10.1177/2515245920918872

107. Palmer W, Okonya O, Jellison S, Horn J, Harter Z, Wilkett M, et al. Intervention reporting of clinical trials published in high-impact cardiology journals: effect of the TIDieR checklist and guide. BMJ Evid-Based Med. 2021;26: 91–97. doi:10.1136/bmjebm-2019-111309

108. Piwowar H A. Who shares? Who doesn’t? Factors associated with openly archiving raw research data. PLoS ONE Electron Resour. 2011;6: e18657. doi:10.1371/journal.pone.0018657

109. Piwowar H A, Chapman W W. Public sharing of research datasets: a pilot study of associations. J Informetr. 2010;4: 148–156.

110. Riedel C, Geßner H, Seegebrecht A, Ayon S I, Chowdhury S H, Engbert R, et al. Including Data Management in Research Culture Increases the Reproducibility of Scientific Results. In: Demmler D, Universitat Hamburg, Vogt-Kolln-Strasse Hamburg, Krupka D, Gesellschaft fur Informatik, Anna-Louisa-Karsch-Strasse Berlin, et al., editors. Gesellschaft fur Informatik (GI); 2022. pp. 1341–1352. doi:10.18420/inf2022\_114

111. Rowhani-Farid A, Aldcroft A, Barnett A G. Did awarding badges increase data sharing in BMJ Open? A randomized controlled trial. R Soc Open Sci. 2020;7: 191818. doi:10.1098/rsos.191818

112. Rowhani-Farid A, Barnett A G. Has open data arrived at the British Medical Journal (BMJ)? An observational study. BMJ Open. 2016;6: e011784. doi:10.1136/bmjopen-2016-011784

113. Rowhani-Farid A, Barnett A G. Badges for sharing data and code at Biostatistics: an observational study. F1000Research. 2018;7: 90. doi:10.12688/f1000research.13477.2

114. Ruamviboonsuk V, Thinkhamrop B, Kulvichit K, Tulvatana W. Data sharing implementation in top 10 ophthalmology journals in 2021. BMJ Open Ophthalmol. 2023;8: 07. doi:10.1136/bmjophth-2023-001276

115. Salameh J P, McInnes M D. F, Moher D, Thombs B D, McGrath T A, Frank R, et al. Completeness of Reporting of Systematic Reviews of Diagnostic Test Accuracy Based on the PRISMA-DTA Reporting Guideline. Clin Chem. 2019;65: 291–301. doi:10.1373/clinchem.2018.292987

116. Savage C J, Vickers A J. Empirical study of data sharing by authors publishing in PLoS journals. PLoS ONE. 2009;4(9) (no pagination): e7078. doi:10.1371/journal.pone.0007078

117. Shamliyan T A, Kane R L. Availability of results from clinical research: failing policy efforts. J Epidemiol Glob Health. 2014;4: 1–12. doi:10.1016/j.jegh.2013.08.002

118. Shen T, Conway C, Rempfert K R, Kyle J E, Colby S M, Gaul D A, et al. The unknown lipids project: harmonized methods improve compound identification and data reproducibility in an inter-laboratory untargeted lipidomics study. BioRxiv Prepr Serv Biol. 2023;03: 03. doi:10.1101/2023.02.01.526566

119. Siebert M, Gaba J F, Caquelin L, Gouraud H, Dupuy A, Moher D, et al. Data-sharing recommendations in biomedical journals and randomised controlled trials: an audit of journals following the ICMJE recommendations. BMJ Open. 2020;10: e038887. doi:10.1136/bmjopen-2020-038887

120. Sims M T, Bowers A M, Fernan J M, Dormire K D, Herrington J M, Vassar M. Trial registration and adherence to reporting guidelines in cardiovascular journals. Heart. 2018;104: 753–759. doi:10.1136/heartjnl-2017-312165

121. Sims M T, Sanchez Z C, Herrington J M, Hensel J B, Henning N M, Scheckel C J, et al. Shoulder Arthroplasty Trials Are Infrequently Registered: A Systematic Review of Trials. PLoS ONE Electron Resour. 2016;11: e0164984. doi:10.1371/journal.pone.0164984

122. Sinaci A A, Gencturk M, Teoman H A, Laleci Erturkmen, G B, Alvarez-Romero C, et al. A Data Transformation Methodology to Create Findable, Accessible, Interoperable, and Reusable Health Data: Software Design, Development, and Evaluation Study. J Med Internet Res. 2023;25: e42822. doi:10.2196/42822

123. Soderberg C K, Errington T M, Schiavone S R, Bottesini J, Thorn F S, Vazire S, et al. Initial evidence of research quality of registered reports compared with the standard publishing model. Nat Hum Behav. 2021;5: 990–997. doi:10.1038/s41562-021-01142-4

124. Speich B, Mann E, Schonenberger C M, Mellor K, Griessbach A N, Dhiman P, et al. Reminding Peer Reviewers of Reporting Guideline Items to Improve Completeness in Published Articles: Primary Results of 2 Randomized Trials. JAMA Netw Open. 2023;6: e2317651. doi:10.1001/jamanetworkopen.2023.17651

125. Struthers C, Harwood J, de Beyer J A, Dhiman P, Logullo P, Schlüssel M. GoodReports: developing a website to help health researchers find and use reporting guidelines. BMC Med Res Methodol. 2021;21: 217. doi:10.1186/s12874-021-01402-x

126. Thelwall M, Kousha K. Do journal data sharing mandates work? Life sciences evidence from Dryad. Aslib J Inf Manag. 2017;69: 36–45. doi:10.1108/ajim-09-2016-0159

127. Toth A A, Banks G C, Mellor D, O’Boyle E H, Dickson A, Davis D J, et al. Study Preregistration: An Evaluation of a Method for Transparent Reporting. J Bus Psychol. 2021;36: 553–571. doi:10.1007/s10869-020-09695-3

128. Trisovic A, Lau M K, Pasquier T, Crosas M. A large-scale study on research code quality and execution. Sci Data. 2022;9: 60. doi:10.1038/s41597-022-01143-6

129. van der Braak K, Ghannad M, Orelio C, Heus P, Damen J A. A, Spijker R, et al. The score after 10 years of registration of systematic review protocols. Syst Rev. 2022;11: 191. doi:10.1186/s13643-022-02053-9

130. Vassar M, Pollard J, Rorah D, Jellison S, Harter Z J, Brasseux S. Assessment of the completeness of intervention reporting of randomized clinical trials for alcohol use disorders: Effect of the TIDieR checklist and guide. Drug Alcohol Depend. 2020;208: 107824. doi:10.1016/j.drugalcdep.2019.107824

131. Veroniki A A, Tsokani S, Zevgiti S, Pagkalidou I, Kontouli K M, Ambarcioglu P, et al. Do reporting guidelines have an impact? Empirical assessment of changes in reporting before and after the PRISMA extension statement for network meta-analysis. Syst Rev. 2021;10: 246. doi:10.1186/s13643-021-01780-9

132. Walsh S, Jones M, Bressington D, McKenna L, Brown E, Terhaag S, et al. Adherence to COREQ Reporting Guidelines for Qualitative Research: A Scientometric Study in Nursing Social Science. Int J Qual Methods. 2020;19. doi:10.1177/1609406920982145

133. Witwer K W. Data submission and quality in microarray-based MicroRNA profiling. Clin Chem. 2013;59(2): 392–400. doi:10.1373/clinchem.2012.193813

134. OECD. Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development. Paris: Organisation for Economic Co-operation and Development; 2015. Available: https://www.oecd-ilibrary.org/science-and-technology/frascati-manual-2015\_9789264239012-en

135. Fletcher RH, Black B. “Spin” in scientific writing: scientific mischief and legal jeopardy. Med Law. 2007;26: 511–525.

136. Rahal R-M, Hamann H, Brohmer H, Pethig F. Sharing the Recipe: Reproducibility and Replicability in Research Across Disciplines. Res Ideas Outcomes. 2022;8: e89980. doi:10.3897/rio.8.e89980

137. Squazzoni F, Ahrweiler P, Barros T, Bianchi F, Birukou A, Blom HJJ, et al. Unlock ways to share data on peer review. Nature. 2020;578: 512–514. doi:10.1038/d41586-020-00500-y

138. Horbach SPJM, Tijdink JK, Bouter L. Research funders should be more transparent: a plea for open applications. R Soc Open Sci. 2022;9: 220750. doi:10.1098/rsos.220750

139. Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull. 1979;86: 638–641. doi:10.1037/0033-2909.86.3.638

140. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias — An Updated Review. PLOS ONE. 2013;8: e66844. doi:10.1371/journal.pone.0066844

141. Song F, Hooper L, Loke YK. Publication bias: what is it? How do we measure it? How do we avoid it? Open Access J Clin Trials. 2013;5: 71–81. doi:10.2147/OAJCT.S34419

142. Coles N. The Red Team Challenge (Part 1): Why I placed a bounty on my own research. In: The 100% CI [Internet]. 29 Jun 2020 [cited 21 May 2024]. Available: https://www.the100.ci/2020/06/29/red-team-part-1/

# Supplementary material

Index

1. Deviations to the original protocol
2. Search strategies
3. Data-extraction items and their answering options
4. List of interventions
5. List of outcomes
6. Explanations of terminology used
7. PRISMA-ScR

# 1.Deviations to the original protocol

Papers detailing barriers and facilitators to promoting reproducibility were extracted during our searches, per protocol, but will be analysed separately in a future manuscript. We also planned to formally apply reference snowballing to retrieve extra papers citing the papers we already included. However, due to the number of articles included this was deemed to be unmanageable, we decided to limit reference checks to relevant reviews identified during screening.

We did not extract the type of publication since the articles were already sorted into categories at the previous step. We did not report the specific name of the design but simplified it by only distinguishing between “comparative”, “post-test design”, “qualitative” and “other” (see figure xx). Also, we added an item to record experimental design characteristics. We decided to not include disclosure of competing interest as proxy since we believe that it is not related closely enough to function as proxy outcome. Also, we excluded computational reproducibility practices from the outcome list since these would be covered by other items such as data sharing and reproducibility checks. We added the term research process transparency for clarification. Furthermore, we excluded assay reproducibility because most of the studies assessing assay reproducibility that we screened did not aim to evaluate reproducibility of a research project.

* Deviations to the screening process: single screening for title/abstract, maybe category to indicate edge cases for a second (senior) reviewer
* Grey literature search not as broad as originally planned?
* Data extraction not in duplicate, but instead checked by another author which gave feedback during extraction, data check in the end to look for missing data and obvious inconsistencies
* No authors contacted is information was insufficient due to time constraints
* Evidence maps were created elsewhere, to retain flexibility. As a consequence, they are not yet interactive.

# 2.Search Strategies

***EMBASE through Ovid: Embase Classic+Embase <1947 to 2023 August 14>***

|  |  |  |
| --- | --- | --- |
| **#** | **Searches** | **Results** |
| 1 | (((data or code or workflow or practices or materials or notebook) adj2 (open or share or shared or sharing or preservation or stewardship)) or "open science" or ((computational or data or open or research or conclusion\* or inferential or analytic or conceptual or direct or exact or statistical) adj3 (reproducib\* or replicability or replicable)) or (research adj5 (transparen\* or credib\*))).ti,ab. or (reporting adj3 guideline\*).ti. | 37043 |
| 2 | Randomized Controlled Trial/ or Controlled Clinical Trial/ or Quasi Experimental Study/ or Pretest Posttest Control Group Design/ or Time Series Analysis/ or Experimental Design/ or Multicenter Study/ or random\*.ti,ab. or groups.ab. or (trial or multicentre or multicenter or multi centre or multi center).ti. or (intervention? or effect? or impact? or controlled or control group? or (before adj12 after) or (pre adj5 post) or ((pretest or pre test) and (posttest or post test)) or quasiexperiment\* or quasi experiment\* or pseudo experiment\* or pseudoexperiment\* or evaluat\* or time series or time point? or repeated measur\* or ((experimental or empirical or qualitative) adj5 (study or studies))).ti,ab. | 18022174 |
| 3 | exp clinical study/ or (Cohort adj (study or studies)).mp. or (Case control adj (study or studies)).tw. or (follow up adj (study or studies)).tw. or (observational adj (study or studies)).tw. or (epidemiologic$ adj (study or studies)).tw. or (cross sectional adj (study or studies)).tw. or (comparative adj stud\*).mp. | 13437868 |
| 4 | 2 or 3 | 24369828 |
| 5 | 1 and 4 | 23517 |
| 6 | limit 5 to embase | 12618 |

***Medline: Ovid MEDLINE(R) ALL <1946 to August 14, 2023>***

|  |  |  |
| --- | --- | --- |
| **#** | **Searches** | **Results** |
| 1 | (((data or code or workflow or practices or materials or notebook) adj2 (open or share or shared or sharing or preservation or stewardship)) or "open science" or ((computational or data or open or research or conclusion\* or inferential or analytic or conceptual or direct or exact or statistical) adj3 (reproducib\* or replicability or replicable)) or (research adj5 (transparen\* or credib\*))).ti,ab. or (reporting adj3 guideline\*).ti. | 29136 |
| 2 | (Clinical study/ or Case control study/ or Family study/ or Longitudinal study/ or Retrospective study/ or Prospective study/ or Cohort analysis/ or Comparative Study/ or (Cohort adj (study or studies)).mp. or (Case control adj (study or studies)).tw. or (follow up adj (study or studies)).tw. or (observational adj (study or studies)).tw. or (epidemiologic$ adj (study or studies)).tw. or (cross sectional adj (study or studies)).tw. or (comparative adj stud\*).mp. or (("randomized controlled trial" or "controlled clinical trial" or "multicenter study" or "pragmatic clinical trial").pt. or non-randomized controlled trials as topic/ or interrupted time series analysis/ or controlled before-after studies/ or random\*.ti,ab. or groups.ab. or (trial or multicenter or "multi center" or multicentre or "multi centre").ti. or (intervention? or effect? or impact? or controlled or control group? or (before adj12 after) or (pre adj5 post) or ((pretest or "pre test") and (posttest or "post test")) or quasiexperiment\* or quasi experiment\* or pseudo-experiment\* or pseudoexperiment\* or evaluat\* or "time series" or time-point? or "repeated measur\*" or ((experimental or empirical or qualitative) adj5 (study or studies))).ti,ab.)) not ((news or comment or editorial).pt. or comment on.cm.) | 15108526 |
| 3 | 1 and 2 | 15811 |

***ERIC <1965 to July 2023>***

|  |  |  |
| --- | --- | --- |
| **#** | **Searches** | **Results** |
| 1 | (((data or code or workflow or practices or materials or notebook) adj2 (open or share or shared or sharing or preservation or stewardship)) or "open science" or ((computational or data or open or research or conclusion\* or inferential or analytic or conceptual or direct or exact or statistical) adj3 (reproducib\* or replicability or replicable)) or (research adj5 (transparen\* or credib\*))).ti,ab. or (reporting adj3 guideline\*).ti. | 3020 |
| 2 | (Clinical study/ or Case control study/ or Family study/ or Longitudinal study/ or Retrospective study/ or Prospective study/ or Cohort analysis/ or Comparative Study/ or (Cohort adj (study or studies)).mp. or (Case control adj (study or studies)).tw. or (follow up adj (study or studies)).tw. or (observational adj (study or studies)).tw. or (epidemiologic$ adj (study or studies)).tw. or (cross sectional adj (study or studies)).tw. or (comparative adj stud\*).mp. or (("randomized controlled trial" or "controlled clinical trial" or "multicenter study" or "pragmatic clinical trial").pt. or non-randomized controlled trials as topic/ or interrupted time series analysis/ or controlled before-after studies/ or random\*.ti,ab. or groups.ab. or (trial or multicenter or "multi center" or multicentre or "multi centre").ti. or (intervention? or effect? or impact? or controlled or control group? or (before adj12 after) or (pre adj5 post) or ((pretest or "pre test") and (posttest or "post test")) or quasiexperiment\* or quasi experiment\* or pseudo-experiment\* or pseudoexperiment\* or evaluat\* or "time series" or time-point? or "repeated measur\*" or ((experimental or empirical or qualitative) adj5 (study or studies))).ti,ab.)) not (news or comment or editorial).pt. | 743872 |
| 3 | 1 and 2 | 1379 |

***APA PsycInfo <1806 to August Week 1 2023>***

|  |  |  |
| --- | --- | --- |
| **#** | **Searches** | **Results** |
| 1 | (((data or code or workflow or practices or materials or notebook) adj2 (open or share or shared or sharing or preservation or stewardship)) or "open science" or ((computational or data or open or research or conclusion\* or inferential or analytic or conceptual or direct or exact or statistical) adj3 (reproducib\* or replicability or replicable)) or (research adj5 (transparen\* or credib\*))).ti,ab. or (reporting adj3 guideline\*).ti. | 7023 |
| 2 | (Clinical study/ or Case control study/ or Family study/ or Longitudinal study/ or Retrospective study/ or Prospective study/ or Cohort analysis/ or Comparative Study/ or (Cohort adj (study or studies)).mp. or (Case control adj (study or studies)).tw. or (follow up adj (study or studies)).tw. or (observational adj (study or studies)).tw. or (epidemiologic$ adj (study or studies)).tw. or (cross sectional adj (study or studies)).tw. or (comparative adj stud\*).mp. or (("randomized controlled trial" or "controlled clinical trial" or "multicenter study" or "pragmatic clinical trial").pt. or non-randomized controlled trials as topic/ or interrupted time series analysis/ or controlled before-after studies/ or random\*.ti,ab. or groups.ab. or (trial or multicenter or "multi center" or multicentre or "multi centre").ti. or (intervention? or effect? or impact? or controlled or control group? or (before adj12 after) or (pre adj5 post) or ((pretest or "pre test") and (posttest or "post test")) or quasiexperiment\* or quasi experiment\* or pseudo-experiment\* or pseudoexperiment\* or evaluat\* or "time series" or time-point? or "repeated measur\*" or ((experimental or empirical or qualitative) adj5 (study or studies))).ti,ab.)) not (news or comment or editorial).pt. | 2710373 |
| 3 | 1 and 2 | 3917 |

***Scopus (August 2023)***

|  |  |  |
| --- | --- | --- |
| 1 | ( TITLE-ABS ( ( ( data OR code OR workflow OR practices OR materials OR notebook ) W/1 ( share OR shared OR sharing OR preservation OR stewardship ) ) OR "open science" OR ( ( computational OR data OR open OR research OR conclusion\* OR inferential OR analytic OR conceptual OR direct OR exact OR statistical ) W/2 ( reproducib\* OR replicability OR replicable ) ) OR ( research W/2 ( transparen\* OR credib\* ) ) ) OR TITLE ( reporting W/2 guideline\* ) ) AND ( TITLE-ABS ( cohort W/1 ( study OR studies ) ) OR TITLE-ABS ( "Case control" W/1 ( study OR studies ) ) OR TITLE-ABS ( follow-up W/1 ( study OR studies ) ) OR TITLE-ABS ( observational W/1 ( study OR studies ) ) OR TITLE-ABS ( "cross sectional" W/1 ( study OR studies ) ) OR TITLE-ABS ( comparative W/1 stud\* ) OR TITLE-ABS ( random\* ) OR TITLE ( multicenter OR "multi center" OR multicentre OR "multi centre" ) OR TITLE-ABS ( intervention\* OR effectiveness OR ( controlled W/3 trial\* ) OR "control group\*" OR ( before W/11 after ) OR ( ( pretest OR "pre test" ) AND ( posttest OR "post test" ) ) OR quasiexperiment\* OR "quasi experiment\*" OR "pseudo experiment\*" OR pseudoexperiment\* OR evaluation OR ( ( experimental OR empirical OR qualitative ) W/4 ( study OR studies ) ) ) ) AND ( EXCLUDE ( DOCTYPE , "cp" ) ) | 12960 |

***World of Science (August 2023)***

|  |  |  |
| --- | --- | --- |
| 1 | ((TI= (((data OR code OR workflow OR practices OR materials OR notebook) NEAR/1 ( share OR shared OR sharing OR preservation OR stewardship)) OR "open science" OR ((computational OR data OR open OR research OR conclusion\* OR inferential OR analytic OR conceptual OR direct OR exact OR statistical) NEAR/2 reproducib\*) OR (research NEAR/2 (transparen\* OR credib\*)) OR (reporting NEAR/2 guideline\*))) AND (TI= (cohort NEAR/1 (study OR studies)) OR TI= ("Case control" NEAR/1 (study OR studies)) OR TI= (follow-up NEAR/1 (study OR studies)) OR TI= (observational NEAR/1 (study OR studies)) OR TI= ("cross sectional" NEAR/1 (study OR studies)) OR TI= (comparative NEAR/1 stud\*) OR TI= (random\*) OR TI= (multicenter OR "multi center" OR multicentre OR "multi centre") OR TI= (intervention\* OR effectiveness OR (controlled NEAR/3 trial\*) OR "control group\*" OR (before NEAR/12 after) OR ((pretest OR "pre test") AND (posttest OR "post test")) OR quasiexperiment\* OR "quasi experiment\*" OR "pseudo experiment\*" OR pseudoexperiment\* OR evaluation OR ((experimental or empirical or qualitative) NEAR/4 (study or studies))))) OR ((AB= (((data OR code OR workflow OR practices OR materials OR notebook) NEAR/1 ( share OR shared OR sharing OR preservation OR stewardship)) OR "open science" OR ((computational OR data OR open OR research OR conclusion\* OR inferential OR analytic OR conceptual OR direct OR exact OR statistical) NEAR/2 reproducib\*) OR (research NEAR/2 (transparen\* OR credib\*)))) AND (AB= (cohort NEAR/1 (study OR studies)) OR AB= ("Case control" NEAR/1 (study OR studies)) OR AB= (follow-up NEAR/1 (study OR studies)) OR AB= (observational NEAR/1 (study OR studies)) OR AB= ("cross sectional" NEAR/1 (study OR studies)) OR AB= (comparative NEAR/1 stud\*) OR AB= (random\*) OR AB= (multicenter OR "multi center" OR multicentre OR "multi centre") OR AB= (intervention\* OR effectiveness OR (controlled NEAR/3 trial\*) OR "control group\*" OR (before NEAR/12 after) OR ((pretest OR "pre test") AND (posttest OR "post test")) OR quasiexperiment\* OR "quasi experiment\*" OR "pseudo experiment\*" OR pseudoexperiment\* OR evaluation OR ((experimental or empirical or qualitative) NEAR/4 (study or studies))))) | 9232 |

# 3.Data-extraction items and their answering options

1. Source, options:
   1. Peer-reviewed
   2. Pre-print
   3. Grey literature
2. Design
   1. Research Question/ Hypothesis (Open text field)
   2. Study design category, options:
      1. Comparative (between-subject comparison)
      2. Comparative (within-subject comparison/repeated measures design)
      3. Posttest design (only a post measurement after the implementation of an intervention and the intervention is explicitly mentioned)
      4. Qualitative
      5. Other
      6. Not applicable
      7. Not reported
   3. Experimental design characteristics, options:
      1. Manipulation
      2. Randomization
      3. Not applicable
      4. Not reported
3. Sample
   1. Sample, options:
      1. Articles/studies
      2. PhD students
      3. Post docs
      4. Academic staff
      5. Reviewers
      6. Academic support staff
      7. Journals
      8. Publishers
      9. Institutions
      10. Funders
      11. Other
4. Sample description:
   1. Country, options:
      1. Country - open text field
      2. Not applicable
      3. Not reported
   2. Broad academic field classification, options:
      1. Natural sciences
      2. Engineering and technology
      3. Medical and health sciences
      4. Agricultural and veterinary sciences
      5. Social sciences
      6. Humanities and the arts
      7. Other
      8. Not applicable
      9. Not reported
   3. Sample size; open text field
5. Intervention
   1. Intervention classification (*each class has several subclasses (see appendix X); multiple options could be chosen*), options:
      1. Open methodology
      2. Open science policies and guidelines
      3. Open data and materials
      4. Open science ethics
      5. Open evaluations
      6. Open educational resources
      7. Open science tools
      8. Other
6. Intervention Details
   1. Level of independent variable (Who)
      1. Government
      2. Funder
      3. Publisher
      4. Journal
      5. Institution
      6. Reviewers
      7. PhD students
      8. Post docs
      9. Academic staff
      10. Academic support staff
      11. Other
      12. Not applicable
   2. Level of dependent variable (Who is affected), options:
      1. Government
      2. Funder
      3. Publisher
      4. Journal
      5. Institution
      6. Reviewers
      7. PhD students
      8. Post docs
      9. Academic staff
      10. Academic support staff
      11. Other
      12. Not applicable
   3. Stage of the process, options:
      1. Before research
      2. During research
      3. After research before or at publication
      4. After publication
      5. Other
   4. Stringency
      1. Mandatory
      2. Optional
      3. Other
      4. Not applicable
      5. Not reported
7. Outcome
   1. Direct outcomes, options:
      1. Inferential reproducibility
      2. Computational/Outcome/Result/Numerical/Statistical reproducibility
      3. Method /methodological reproducibility
      4. Inferential replicability
      5. Computational/Outcome/Result/Numerical/Statistical replicability
      6. Other
   2. Outcome Details N/A
      1. Outcome metric/operationalization: open text field
      2. When was outcome measured (in relation to intervention): open text field
8. Proxy outcomes (*each class has several subclasses (see appendix X); multiple options could be chosen*); options:
   1. Registration Status
   2. Review
   3. Methods transparency
   4. Data
   5. Research material sharing
9. Results
   1. Answer to RQ/Hypothesis: open text field
   2. Relevant descriptives: open text field
   3. Primary statistical/meta analytic or qualitative results for each hypothesis: open text field
   4. Additional results: open text field
10. Important limitations: tick box plus open text field
11. Exclude Out of scope: tick box
12. Data extraction finished: tick box

# 4.List of interventions

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Intervention class** (as per data-extraction file) | **Intervention subclass** (as per data-extraction file) | **n** | **reporting\_sharing** (additional variable) | **intervention\_exact** (additional variable for more details) | **References**  (First author (Year)) |
| **Open methodology** | Methods transparency, open notebook, research transparency | 0 |  |  |  |
| Open protocol | 1 | preregistration | publication of protocol | Gkiouras (2022) |
| Pre-registration, study registration, analysis plan | 5 | preregistration | preregistration | Toth (2021; 3 studies), van der Braak (2022), Shamliyan (2014) |
| Pre-data-collection registration | 0 |  |  |  |
| Registered reports | 2 | preregistration | registered reports | Obels (2020), Soderberg (2021) |
| Analysis script sharing | 0 |  |  |  |
| Analysis transparency | 0 |  |  |  |
| Design transparency | 0 |  |  |  |
| Reproducible research practices | 0 |  |  |  |
| Project workflow | 2 | standardization | standardization of processes | Arroyo-Araujo (2019), Shen (2023) |
| Other | 0 |  |  |  |
| **Open science policies and guidelines** | Rewards and incentives | 9 | data/code sharing | badge for code sharing | Rowhani-Farid (2018; 2 studies) |
| data/code sharing | badge for open data | Crüwell (2023), Hardwicke (2021), Rowhani-Farid (2020) |
| data/code sharing | badge for open data and materials | Kidwell (2016) |
| data/code sharing | badge for open science | Nuijten (2017) |
| open science | badge for open science | Giofre (2017), Giofre (2022) |
| Concordats | 0 |  |  |  |
| Promotion and tenure policies | 0 |  |  |  |
| Reporting guidelines, reporting standards | 15 | reporting guidelines | use of reporting checklist | Caron (2017), Fuller (2014), Nguyen (2022) |
| reporting guidelines | publication of reporting checklist | Agha (2017), Can (2011), Chow (2018), Clayson (2019), Howell (2015), Liu (2022), Meredith (2019), Palmer (2021), Salameh (2019), Vassar (2020), Veroniki (2021) |
| methodological guidelines | publication of methodological guidelines | Fernandes (2019) |
| TOP Guidelines (Transparency and Openness Promotion) | 0 |  |  |  |
| Policy guidelines (e.g., of funders/publishers) | 49 | data/code sharing | journal submission guidelines | Alsheikh-Ali (2011), Hamilton (2022), Hamilton (2023) |
| data/code sharing | journal data/code sharing policy | Askarov (2022), Auspurg (2020) |
| data/code sharing | ICMJE data sharing statement | Bergeat (2022), Danchev (2021), Ruamviboonsuk (2023), Siebert (2020) |
| data/code sharing | journal code sharing policy | Culina (2020) |
| data/code sharing | journal data sharing policy | Federer (2018), Gorman (2020), Magee (2014), Milia (2012), Naudet (2018, 2 studies), Nguyen (2022), Nuijten (2017; 2 studies), Piwowar (2011), Rowhani-Farid (2016), Savage (2009), Thelwall (2017) |
| data/code sharing | journal characteristics | Hamilton (2022) |
| data/code sharing | journal open data policy | Hardwicke (2018), Laurinavichyute (2022; 2 studies) |
| data/code sharing | funder data management policy | Narang (2023) |
| data/code sharing | journal and funder policies | Piwowar (2010) |
| data/code sharing | funder data sharing policy | Piwowar (2011) |
| data/code sharing | journal policy strictness | Trisovic (2022) |
| methodological guidelines | journal submission guidelines | Giofre (2017), Giofre (2022) |
| preregistration | journal accepts regardless outcome | Blanco-Perez (2020) |
| preregistration | journal policy on trial registration | Hardt (2013), McGee (2011), Sims (2018) |
| preregistration | ICMJE proposed trial registration | Hardt (2013), Sims (2016) |
| preregistration | journal submission guidelines | Ladd (2023) |
| preregistration | FDA trial registration policy | Marin (2015) |
| reporting guidelines | journal endorses reporting checklist | Agha (2016), Caron (2017), Hopewell (2012), Walsh (2020) |
| reporting guidelines | ICMJE endorses reporting guideline | Audet (2017) |
| reporting guidelines | journal requires reporting checklist | Goldacre (2019), Witwer (2013) |
| reporting guidelines | journal submission guidelines | Ladd (2023) |
| FAIR Principles | 0 |  |  |  |
| Data access policies | 0 |  |  |  |
| Open science plans, open science university plans | 0 |  |  |  |
| Reproducibility checklists | 0 |  |  |  |
| Citation standards | 0 |  |  |  |
| Other | 0 |  |  |  |
| **Open data and materials** | Data sharing | 0 |  |  |  |
| Data availability | 1 | data/code sharing | authors shared data | Bergeat (2022) |
| Data transparency | 0 |  |  |  |
| Data publishing, data paper, data journals | 0 |  |  |  |
| Open access | 0 |  |  |  |
| Early access | 0 |  |  |  |
| Code sharing | 0 |  |  |  |
| Open data | 0 |  |  |  |
| Open material/s | 0 |  |  |  |
| Open source software, open source, open software and code | 0 |  |  |  |
| Open hardware | 0 |  |  |  |
| Open experiment scripts | 0 |  |  |  |
| Pre-prints | 0 |  |  |  |
| Presenting metadata clearly | 0 |  |  |  |
| Affirmative sharing declarations | 1 | data/code sharing | data sharing statements | Krawczyk (2012) |
| Data management | 0 |  |  |  |
| Dynamic documents | 0 |  |  |  |
| Other | 0 |  |  |  |
| **Open science ethics** | Competing interest statements | 0 |  |  |  |
| Fundings statements | 0 |  |  |  |
| Retractions | 0 |  |  |  |
| Other | 0 |  |  |  |
| **Open evaluations** | Open peer review - Open identities | 0 |  |  |  |
| Open peer review - Open reports | 0 |  |  |  |
| Open peer review - Open pre-review | 0 |  |  |  |
| Open peer review - Final version commenting | 0 |  |  |  |
| Open metrics | 0 |  |  |  |
| Other | 0 |  |  |  |
| **Open educational resources** | Open educational resources | 0 |  |  |  |
| Open science communities, networks, journal clubs | 0 |  |  |  |
| Training | 2 | standardization | training | Deardorff (2020) |
| systematic reviews | training | Menon (2021) |
| Other | 1 | preregistration | results-free review | Findley (2016) |
| **Open science tools** | Open repositories, data repositories | 0 |  |  |  |
| Open workflow tools, workflow management systems | 1 | standardization | tools | Madduri (2019) |
| Version control | 0 |  |  |  |
| Collaborative platforms | 0 |  |  |  |
| Dockerization, docker | 1 | statistical tool | statistical tool | Chan (2023) |
| Software containers (general frameworks for handling dependencies) | 0 |  |  |  |
| Software versioning | 0 |  |  |  |
| Data presentation software | 0 |  |  |  |
| Continuous integration | 0 |  |  |  |
| Executable Research Compendium | 0 |  |  |  |
| Other | 5 | automatic code cleaning | automatic code cleaning | Trisovic (2022) |
| standardization | standardization of processes | Sinaci (2023) |
| reporting guidelines | web-based writing aid | Barnes (2015), Hopewell (2016), Struthers (2021) |
| **Other** | Type1/Type 2 error reduction | 3 | statistical interventions | effect-size filter | Gregori (2013) |
| statistical interventions | p-value calibration | Gruber (2016; 2 studies) |
| Collaborative research | 1 | standardization | building a (training) community | Riedel (2022) |
| Other | 5 | reporting guidelines | editor checks reporting | Blanco (2020) |
| reporting guidelines | peer review checks reporting | Cobo (2011) |
| reporting guidelines | reminders for peer-reviewers | Speich (2023) |
| open science | role model in open science | Haven (2023) |
| statistical interventions | statistical approach | Jaljuli (2023) |
|  | Total: | 104 |  | Total: |  |

# 5.List of Outcomes

## 5a.Direct outcomes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Outcomes class** (Additional variable) | **Outcomes subclass** (as per data-extraction sheet) | **n** | **Outcomes subclass** (modified for analysis purposes) | **Exact outcomes** (more details) | **References** (First author (year)) |
| Reproducibility | Inferential reproducibility | 1 | Inferential reproducibility | inferential reproducibility | Bergeat (2022) |
| Computational/Outcome/Result/Numerical/Statistical reproducibility | 1 | Results reproducibility | code reproducibility | Chan (2023) |
| 7 | Results reproducibility | results reproducibility | Crüwell (2023), Hardwicke (2021), Laurinavichyute (2022), Madduri (2019), Naudet (2018), Obels (2020), Riedel (2022) |
| Method /methodological reproducibility | 0 |  |  |  |
| Replicability | Inferential replicability | 0 |  |  |  |
| Computational/Outcome/Result/Numerical/Statistical replicability | 2 | Results replicability | results replicability | Arroyo-Araujo (2019), Shen (2023) |
| Other direct outcome | Other | 3 | Statistical inconsistencies | statistical inconsistencies | Nuijten (2017; 3 studies) |
| 1 | Type-I error reduction | type-I error reduction | Jaljuli (2023) |

## 5b.Proxy outcomes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Outcomes class** (Additional variable) | **Outcomes subclass** (as per data-extraction sheet) | **n** | **Outcomes subclass** (modified for analysis purposes) | **Exact outcomes** (more details) | **References** (First author (year)) |
| Registration Status | Pre-registration | 4 | Pre-registration | preregistration | Hardt (2013), McGee (2011), Sims (2016), van der Braak (2022) |
| Registered reports | 0 |  |  |  |
| Review | Post study peer review | 0 |  |  |  |
| Reproducibility checks | 0 |  |  |  |
| Methods transparency | Analysis transparency | 1 | Transparency of Analysis and Design | completeness of reporting | Clayson (2019) |
| Design transparency | 1 | Design transparency | reporting specific elements | Toth (2021) |
| Research process/project workflow transparency | 1 | Research process or workflow transparency | self-reported transparency | Menon (2021) |
| Methods transparency general | 32 | Methods transparency general | completeness of reporting | Agha (2016), Agha (2017), Barnes (2015), Blanco (2020), Can (2011), Caron (2017; 2 studies),Chow (2018), Cobo (2011), Fernandes (2019), Fuller (2014), Giofre (2017), Giofre (2022), Goldacre (2019), Hopewell (2012), Hopewell (2016), Howell (2015), Ladd (2023), Liu (2022), Meredith (2019), Nguyen (2022), Palmer (2021), Salameh (2019), Speich (2023), Struthers (2021), Vassar (2020), Veroniki (2021), Walsh (2020) |
| reporting specific elements | Audet (2017), Ladd (2023), Sims (2018), Toth (2021) |
| Data | Data management plan | 0 |  |  |  |
| Research material sharing | Material sharing | 1 | Data and material sharing | data-availability | Kidwell (2016) |
| Data sharing | 1 | Data sharing and methods transparency | data-availability | Witwer (2013) |
| 10 | Data sharing | data-availability | Hamilton (2023), Alsheikh-Ali (2011), Bergeat (2022), Giofre (2017), Giofre (2022), Gorman (2020), Magee (2014), Piwowar (2010), Piwowar (2011; 2 studies) |
| 12 | data-sharing | Danchev (2021), Hamilton (2022), Haven (2023), Milia (2012), Naudet (2018), Nguyen (2022), Rowhani-Farid (2016; 3 studies), Ruamviboonsuk (2023), Savage (2009), Thelwall (2017) |
| 4 | data-sharing statements | Federer (2018), Hardwicke (2018), Narang (2023), Siebert (2020) |
| 2 | Data and code sharing | data-availability | Auspurg (2020), Laurinavichyute (2022) |
| Code sharing | 3 | Code sharing | code-sharing | Culina (2020), Hamilton (2022), Rowhani-Farid (2018) |
| Software sharing | 0 |  |  |  |
| Research material sharing general | 1 | Research material sharing general | materials sharing | Krawczyk (2012) |
| Miscellaneous | Other | 16 | Other | code execution rate | Trisovic (2022; 2 studies) |
| FAIRness of data | Sinaci (2023) |
| lessons learned | Findley (2016) |
| overall quality | Soderberg (2021) |
| self-reported behaviour | Deardorff (2020) |
| statistical significance | Askarov (2022), Blanco-Perez (2020), Gkiouras (2022), Toth (2021) |
| trial results published | Hardt (2013), Marin (2015), Shamliyan (2014) |
| type I errors | Gregori (2013), Gruber (2016) |
| type II errors | Gruber (2016) |

# 6. Explanations of terminology used

|  |  |
| --- | --- |
| **Term** | **Explanation** |
| Analysis reproducibility | Analysis should lead to the same interpretation (European Commission 2022) |
| Assay reproducibility | Obtain consistent results when testing identical samples in multiple laboratories using the same methods, reagents and controls (Waugh & Clark 2021) |
| Code execution rate | The proportion of executed computer code (e.g., R scripts) without an error (Trisovic *et al.* 2022) |
| Completeness of reporting | Reporting all elements that are required according to the reporting guidelines or standards |
| Computational reproducibility | Obtain the same results using the same input data, computational methods and conditions of analysis (Parsons *et al.* 2022) |
| Conceptual replicability | Evaluate the validity of an effect using a different set of experimental conditions, sample and methods (Parsons *et al*. 2022) |
| Data availability | The accessibility and actual availability of the data |
| Data management plan | Plan describing how research data are collected, generated, processed, analyzed and managed |
| Data sharing statements | A statement in the research report that data are available |
| Direct replicability | Replicability (see *Replicability* entry) assessed as a direct outcome of an open science (or other) intervention |
| Direct reproducibility | Reproducibility (see *Reproducibility* entry) assessed as a direct outcome of an open science (or other) intervention |
| FAIRness of data | Data feature that refers to them as Findable, Accessible, Interoperable and Reusable (Sinaci *et al.* 2023) |
| FAIR principles | Principles to make scholarly material Findable, Accessible, Interoperable and Reusable (FAIR) (Parsons *et al.* 2022) |
| Inferential reproducibility | Draw similar conclusions from a reanalysis or replication of a study (Goodman *et al* 2016) |
| Lessons learned | Qualitative account of conclusions after an experience or experiment with the purpose of further guidance (Findley *et al.* 2016) |
| Methodological reproducibility, method reproducibility | Have enough detail about study procedures and data so the same procedures can be repeated (Goodman *et al.* 2016) |
| Methods transparency   * Analysis transparency * Design transparency * Project workflow transparency | Transparent reporting of the methodology in such detail that methods reproducibility is considered feasible, e.g. by adhering to reporting guidelines |
| Outcome reproducibility | Ability to produce the same outcome as the original experiment (European Commission 2022) |
| Overall paper quality | Subjective judgment about the overall quality of a research paper (Soderberg *et al.* 2021) |
| Pre-registration | Publishing the plan for a study, including research questions and hypotheses, research design and data analysis, before the data have been collected or examined (Parsons *et al.* 2022) |
| Process reproducibility | Repeat the original analysis using the same data and code (Nosek *et al.* 2022) |
| Registered reports | Publication and peer review of study protocol and analysis plan for in principle acceptance before data collection; after data analyses and write-up of results and discussion, stage 2 review assesses whether authors sufficiently followed their study plan and reported deviations from it. This shifts the focus of the review to the study’s research question and methodology and away from the perceived interest in the study’s results. (Parsons *et al*. 2022) |
| Replicability | Obtain the same results for the same research question, using the same analytical method but on a different sample (European Commission 2020) |
| Reporting specific elements | Reporting specific elements that are required according to the reporting guidelines or standards |
| Reproducibility | Obtain the same results when re-enacting a study using the same methodology, data and code (European Commission 2020) |
| Reproducibility checks | Independent checking whether code or other parts of research are reproducible, e.g. by reviewers |
| Research material sharing   * Material sharing * Data sharing * Code sharing * Software sharing | Sharing of all information, tools, equipment etc. that are necessary to reproduce/replicate a study. This includes descriptive metadata sharing and code explanations |
| Results reproducibility | Obtaining the same results from an independent study whose procedures are as closely matched to the original experiment as possible (Goodman *et al.* 2016) |
| Self-reported behaviour | Behaviour changes related to open science practices and data sharing, resulting from programming workshops about researcher workflows, as reported by the interviewees who attended the workshops (Deardorff 2020) |
| Self-reported transparency | Transparency as evaluated and reported by the researchers themselves |
| Statistical inconsistencies | Numerical results that do not match the description of the statistical analysis or that do not match other numerical results mentioned in the paper (Nuijten *et al.* 2017) |
| Statistical reproducibility | Detailed information is provided about the choice of statistical tests, model parameters, and threshold values (The Turing Way Community 2022) |
| Statistical significance | A statistical criterion that refers to the rejection of the null hypothesis at a prespecified significance level (e.g., 1%) (Askarov *et al.* 2022; Blanco-Perez *et al.* 2020; Gkiouras *et al.* 2022; Toth *et al.* 2021) |
| TOP guidelines | Transparency and Openness Promotion Guidelines for published research (Nosek *et al.* 2015) |
| Trial results published | When the scientific results of a clinical trial are published in the scientific literature after an initial preregistration or protocol (Hardt *et al.* 2013; Marin *et al.* 2015; Shamliyan *et al.* 2014) |
| Type-I error | A decision to reject a true null hypothesis (Gregori *et al.* 2013; Gruber *et al.* 2016) |
| Type-I error reduction | A reduction of statistically significant results while the null hypothesis is true (Jaljuli *et al.* 2023) |
| Type-II error | A decision to accept a false null hypothesis (Gruber *et al.* 2016) |

European Commission, Directorate-General for Research and Innovation; Baker L, et al.: Reproducibility of scientific results in the EU: scoping report. Lusoli, W. (editor): Publications Office of the European Union,2020. 10.2777/341654

European Commission, Directorate-General for Research and Innovation: Assessing the reproducibility of research results in EU Framework Programmes for Research: final report. Publications Office of the European Union,2022. 10.2777/186782

Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? *Sci Transl Med*. 2016 Jun 1;8(341):341ps12. 10.1126/scitranslmed.aaf5027. PMID: 27252173

Nosek BA, Alter G, Banks GC, et al. Promoting an open research culture. *Science. 2015;*348:1422-1425 10.1126/science.aab2374

Nosek BA, Hardwicke TE, Moshontz H, et al. Replicability, robustness, and reproducibility in psychological science. *Annu Rev Psychol.* 2022;73:719–748. 34665669 10.1146/annurev-psych-020821-114157

Parsons S, Azevedo F, Elsherif MM et al. A Community-Sourced Glossary of Open Scholarship Terms. *Nature human behaviour*;2022;6(3):312-318. 10.1038/s41562-021-01269-4

The Turing Way Community. The Turing Way: A handbook for reproducible, ethical and collaborative research. 2022. Zenodo. 10.5281/zenodo.3233853.

Waugh C & Clark G. Factors affecting test reproducibility among laboratories. *Revue Scientifique et Technique de l’Office International des Epizooties; 2021;*40(1):31-143. 10.20506/RST.40.1.3213

# 7.Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist

|  |  |  |  |
| --- | --- | --- | --- |
| **Section** |  | **PRISMA-ScR checklist item** | **Reported on page #** |
| **Title** | | | |
| Title | 1 | Identify the report as a scoping review. | Page 1 and 2 |
| **Abstract** | | | |
| Structured summary | 2 | Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives. | Page 2, though unstructured |
| **Iintroduction** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach. | Page 4, last paragraph of introduction |
| Objectives | 4 | Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives. | Page 5 |
| **Methods** | | | |
| Protocol and registration | 5 | Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number. | Page 5; first section of Methods |
| Eligibility criteria | 6 | Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale. | Page 5 to 7 |
| Information sources\* | 7 | Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed. | Page 7 |
| Search | 8 | Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated. | Supplemental file 2 |
| Selection of sources of evidence† | 9 | State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review. | Page 8 |
| Data charting process‡ | 10 | Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators. | Page 8 and 9 |
| Data items | 11 | List and define all variables for which data were sought and any assumptions and simplifications made. | Supplemental file 3 |
| Critical appraisal of individual sources of evidence§ | 12 | If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate). | N/A |
| Synthesis of results | 13 | Describe the methods of handling and summarizing the data that were charted. | Page 9 |
| **Results** | | | |
| Selection of sources of evidence | 14 | Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram. | Page 10 |
| Characteristics of sources of evidence | 15 | For each source of evidence, present characteristics for which data were charted and provide the citations. | Page 10 and 11; Table 1 |
| Critical appraisal within sources of evidence | 16 | If done, present data on critical appraisal of included sources of evidence (see item 12). | N/A |
| Results of individual sources of evidence | 17 | For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives. | Supplementary file 4 and 5 |
| Synthesis of results | 18 | Summarize and/or present the charting results as they relate to the review questions and objectives. | Page 20 |
| **Discussion** | | | |
| Summary of evidence | 19 | Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups. | Page 21; start of Discussion |
| Limitations | 20 | Discuss the limitations of the scoping review process. | Page 23; separate heading |
| Conclusions | 21 | Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps. | Page 24 |
| **Funding** | | | |
| Funding | 22 | Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review. | Page 25 |

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

\* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O’Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting*.*

§The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

*From:* Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Ann Intern Med. 2018;169:467–473. [doi: 10.7326/M18-0850](http://annals.org/aim/fullarticle/2700389/prisma-extension-scoping-reviews-prisma-scr-checklist-explanation).