

# GoodReports - developing a website to help health researchers find and use reporting guidelines

Caroline Struthers (✉ [caroline.struthers@csm.ox.ac.uk](mailto:caroline.struthers@csm.ox.ac.uk))

NDORMS, University of Oxford

James Harwood

NDORMS, University of Oxford

Jennifer Anne de Beyer

NDORMS, University of Oxford

Paula Dhiman

NDORMS, University of Oxford

Patricia Logullo

NDORMS, University of Oxford

Michael Schlüssel

NDORMS, University of Oxford

---

## Research Article

**Keywords:** Standards, software, reporting guidelines, education, reproducibility

**Posted Date:** March 15th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-289510/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The EQUATOR Network improves the quality and transparency in health research, primarily by promoting awareness and use of reporting guidelines. In 2018, the UK EQUATOR Centre launched GoodReports.org, a website that helps authors find and use reporting guidelines. This paper describes the tool's development so far. We evaluated user experience and behaviour while using the website as part of manuscript submission to a journal to inform future development.

**Methods:** We conducted a survey to collect data on users' experience of the GoodReports website during manuscript submission. We assessed the tool's reliability by checking our agreement with the tool's checklist recommendation on a random sample of manuscripts submitted to a partner journal. We compared the proportion of authors submitting a reporting checklist alongside their manuscripts between groups exposed or not exposed to the GoodReports tool. We compared the text of manuscripts before an author received a reporting guideline recommendation with the text subsequently submitted to the partner journal.

**Results:** Seventy percent (423/599) of survey respondents rated GoodReports 8 or more out of 10 for usefulness, and 74% (198/267) said they had made changes to their manuscript after using the website. We agreed with the GoodReports reporting guideline recommendation in 84% (72/86) of cases. Of authors who completed the guideline finder questionnaire, 14% (10/69) failed to submit a completed checklist compared to 30% (41/136) who did not use the tool. Of the 69 authors who received a GoodReports reporting guideline recommendation, 20 manuscript pairs were included in a before-and-after study. Five included more information in their methods section after exposure to GoodReports. On average, authors reported 57% of necessary reporting items before completing a checklist on GoodReports.org and 60% after.

**Conclusion:** The data provide encouraging signs that GoodReports could increase the use of reporting guidelines. They also underline the need for reporting guidance to be introduced early in the writing process. We are developing GoodReports by adding more reporting guidelines to the database, and by developing the functionality to integrate reporting items into Word article templates. We will test whether GoodReports users write more complete study reports in a randomised trial.

## Background

Around 80% of articles reporting health-related research do not include enough detail for a reader to fully understand, assess, and replicate the methods and results (1). Reporting guidelines aim to solve this problem by specifying the minimum information authors need to include when writing up their research for publication. Reporting guideline documents often include a checklist, which many medical journals ask authors to submit alongside their manuscripts as evidence they have included all of the necessary information.

The EQUATOR (Enhancing Quality and Transparency of Health Research) Network is an international initiative that has been providing support for the dissemination and use of reporting guidelines since 2006. The UK EQUATOR Centre makes reporting guidelines more accessible by maintaining a centralised, searchable database alongside resources and training to support their use (2). Despite the work of EQUATOR and many other organisations, such as NC3Rs, Cochrane, ICMJE, WAME, EASE, and COPE (3-8), to promote the use of reporting guidelines, reporting quality remains poor (9) and use of even the most popular guidelines remains low (10, 11).

A scoping review of interventions to improve adherence to reporting guidelines found a lack of practical training on how to use them and that guidelines were not easy to access or understand (12). There are over 400 reporting guidelines in the EQUATOR database, which differ in the amounts of instruction they provide. Authors may struggle to choose an appropriate guideline. Many are published behind paywalls and in unusable formats, such as PDF checklists that cannot be filled in.

To address these issues, the UK EQUATOR Centre has created GoodReports.org (13), a website that helps authors select the most appropriate reporting guideline for their study and gives them immediate access to a user-friendly checklist. Authors can fill in the checklist online and download it to include with their journal submission.

This paper describes our development and testing of GoodReports.org so far. The UK EQUATOR Centre has partnered with Penelope.ai (14), a software company already providing a manuscript-checking service to *BMJ Open*. This partnership allowed us to quickly drive traffic to the GoodReports website and observe whether authors who completed a GoodReports checklist and submitted it to the journal with their article added information to their manuscripts. We also gathered qualitative user feedback from authors using Penelope to assess whether this workflow to access reporting checklists is acceptable and how future development of GoodReports could better serve author needs.

## Methods

### Developing GoodReports.org

The GoodReports website has two main features: authors can complete a questionnaire about their study to receive a reporting guideline suggestion, then can immediately access reporting checklists to fill out on- or offline. Each checklist includes clear instructions, and each reporting item is linked to an explanation of why that item is needed and examples of good reporting, whenever guideline developers have provided such information.

We decided to include reporting guidelines that cover the main generic study designs and are commonly recommended by journals. We started with the 13 popular reporting guidelines highlighted on the EQUATOR homepage. Although published as one reporting guideline, STROBE covers three observational study designs and is made available by the STROBE development group as three separate checklists. We

included all three. We added STREGA for genetic association studies (26) as it is included in the *BMJ Open* guide for authors. The 16 reporting guidelines included in GoodReports are shown in Table 1.

**Table 1: List of reporting guidelines in the GoodReports database (15-28)**

	Name	Study type
1	ARRIVE	Laboratory animals*
2	CARE	Case reports
3	CHEERS	Economic evaluations
4	CONSORT	Randomised trials*
5	MOOSE	Meta-analyses of observational studies
6	PRISMA	Systematic reviews and meta-analyses*
7	PRISMA-P	Protocols of systematic reviews
8	SPIRIT	Protocols of randomised trials*
9	SQUIRE	Quality improvement studies
10	SRQR	Qualitative studies
11	STARD	Diagnostic test accuracy*
12	STREGA	Genetic association studies
13	STROBE case control	Case-control studies*
14	STROBE cohort	Cohort studies*
15	STROBE cross sectional	Cross-sectional studies*
16	TRIPOD	Prognostic studies
* GoodReports entry includes link to explanation and examples		

We adapted the original questionnaire to our new set of guidelines (Figure 2).

We reduced the use of research and methods jargon as much as possible to improve accessibility and clarified some of the questions in response to initial user feedback. We used multiple-choice options to keep the decision tree short and easy to navigate. The publishers' copyright licence for MOOSE (19) and SRQR (24) did not automatically allow us to reuse the content to create an openly accessible online checklist. We were granted permission to do so on payment of a licence fee for one and two years respectively, which covered the duration of this study.

## Reaching users

GoodReports.org went live in January 2018. We used Penelope.ai (14), a company owned by co-author JH, to attract users. Penelope.ai provides software to journals that automatically checks new submissions and gives immediate feedback to authors to help them meet journal requirements. Penelope.ai allowed us to collaborate with *BMJ Open*, one of Penelope.ai's customers, to capture authors in the process of submitting their articles for publication. All authors submitting to *BMJ Open* can opt to use Penelope.ai, but it is not mandatory.

From January 2018, the Penelope.ai upload form was amended to include the GoodReports guideline recommendation questionnaire (Figure 2). Authors who used Penelope.ai therefore had to answer the checklist finder questionnaire before uploading their manuscript. When appropriate, Penelope.ai's feedback report on their manuscript included a recommendation to use a reporting guideline and a link to the associated checklist on GoodReports.org.

Below is a representative author journey from one of Penelope.ai's client journals, *BMJ Open*.

1. Authors begin their submission on BMJ Open: <https://bmjopen.bmj.com/pages/authors/>
2. Authors receive an option for an automated manuscript check: <https://app.penelope.ai/manuscript-check/q/bmjopen>
3. Authors that opt for an automated check answer a few questions about their work and view their feedback online: <https://app.penelope.ai/submissions/demo/>
4. Depending on the information authors give when uploading to Penelope.ai, their feedback may include an instruction to complete a reporting guideline on <https://www.goodreports.org>

## 1. Individual user feedback

From 25 January 2018 to 6 November 2019, all Penelope.ai users who received a reporting checklist recommendation as part of their manuscript report were sent an automated email survey a day later about their experience of using GoodReports.org. The sample size was determined by the number of users within that timeframe. We used Typeform.com (40) to collect the survey data, with the first question embedded in the email. We asked users:

1. "You were recently advised to complete a checklist at [www.goodreports.org](http://www.goodreports.org). How useful did you find the checklist?" (rating scale: 0 (least useful) to 10)

1. If a rating of 7 or lower: "Can you explain why you gave a rating of [number]?" (multiple choice. We set 7 as the cut-off for this question as we anticipated that our median rating would be 8/10)

- The checklist items were not relevant to my work
- The checklist was too long
- The checklist was confusing

- The website was confusing
- Other (free text)

2. "How could we make [www.goodreports.org](http://www.goodreports.org) more useful?" (free text)

3. From December 2018: "After using the checklist did you make any changes to your manuscript?" (yes/no)

1. If yes, "What did you change?" (free text)
2. If no, "Why didn't you make any changes?" (free text)

Quantitative responses are reported as counts. The free-text responses for question 3.1 and 3.2 were read and discussed by authors CS and JH. JH then extracted general themes. We report the themes, their frequency, and representative quotes.

## **2. Reliability of the questionnaire in helping authors find the most appropriate reporting guideline for their work**

To determine whether the checklist finder questionnaire generally led users to an appropriate checklist, we selected 100 manuscripts from all of those uploaded to Penelope.ai by *BMJ Open* authors between 25 January 2018 and 16 February 2019. Manuscripts were randomly selected from Penelope.ai's database using Python's `randint` function. The sample size was determined by the amount of time available for the two investigators to assess the sample of manuscripts.

CS and JH separately read the titles, abstracts, and methods section of each manuscript and decided which, if any, of the 16 guidelines in the GoodReports database should have been recommended. They compared recommendations and resolved conflicts through discussion. They were blinded to the checklist finder recommendation up to this point. The final assessor recommendation was then compared with the recommendation that authors received from the checklist finder questionnaire. Where the recommendations differed, the assessors examined the questionnaire responses and the corresponding manuscripts together and judged the most likely reason for the discrepancy.

We report the percentage of manuscripts where the checklist finder questionnaire and assessor recommendations matched and possible reasons for mismatches. No statistical tests were done, as the purpose was not to prove that the checklist finder questions lead to appropriate checklists, but to identify how the questionnaire could be improved before a more rigorous evaluation.

### **Use of GoodReports, checklist submission rates, and manuscript completeness**

*BMJ Open* is an existing customer of Penelope.ai. The Editor-in-Chief agreed to allow us access to submitted manuscripts to gather initial data on whether directing users to GoodReports at the point of submission was useful and for use in guiding further development.

We used a sample of *BMJ Open* submissions to observe how exposure to GoodReports correlated with submission quality. We were interested in whether authors included reporting checklists in their submission and whether using a reporting checklist led authors to add missing information to their manuscripts.

We collected data from all newly submitted manuscripts checked by *BMJ Open* staff on 9, 10, 11, 25, 28, and 29 May 2018. These dates were selected by the journal and determined the sample size. It was not practical for the journal to increase the length of the data collection period. We only included manuscripts checked for the first time on these dates. We excluded manuscripts that had been first submitted outside the recording window, returned to the author for corrections, and resubmitted within the recording window.

### **3. Exposure to GoodReports.org and rates of submission of a completed reporting checklist**

*BMJ Open* shared some of the data they collect in their normal day-to-day activity with us, such as whether the submission had previously been checked and notes from the technical editor about unmet journal requirements. These data did not specifically include whether the author had included a reporting checklist in their submission. However, the journal enforces checklists, and we could see the editor's notes. We were therefore able to count the number of "checklists noted to be missing."

We split submissions into two groups, those whose authors had opted to check their manuscript with Penelope.ai before submission and received a checklist recommendation, when appropriate, and those whose authors had opted not to use the checker. JH identified whether an author had used Penelope.ai by searching the Penelope.ai logs for the author's email address and cross-referencing the manuscript's file names and titles, without knowing whether a reporting guideline had been submitted for that manuscript.

We report the proportion of manuscripts where a checklist had been flagged as missing for each group.

### **4. Completeness of reporting before and after using a GoodReports reporting checklist**

We observed whether authors that used and submitted a reporting guideline checklist from GoodReports.org changed their manuscript and improved the completeness of their reporting as a result.

We started with the subset of manuscripts from the study on submission rates that had:

- 1) been checked by Penelope.ai before submission to *BMJ Open*,
- 2) not withdrawn their submission from *BMJ Open*, and
- 3) included a reporting guideline checklist from GoodReports.org.

We conducted a before-and-after study on the included manuscripts. The version that was submitted to Penelope.ai for an automatic pre-submission check (the "before" version) was compared to the version subsequently submitted to *BMJ Open* (the "after" version).

Manuscripts submitted with checklists obtained elsewhere, such as the EQUATOR Network website or the journal website, were excluded. We wanted to reduce the chance that the authors of manuscripts in our “before” group had used a checklist before visiting GoodReports.org.

JH redacted the title and methods sections of the “before” and “after” versions so that no personal information was shared with assessors. The “before” versions were all in .docx format, so text could be copied and pasted into a fresh Microsoft Word file. The “after” versions were PDFs as *BMJ Open* automatically converts submissions into PDF and adds watermarks, line numbers, and footers. JH split PDF files into smaller files containing only the title and methods sections for data extraction. These differing file formats meant that assessors could not be blinded as to whether the manuscript was the “before” or “after” version.

Five assessors (JdB, MS, PD, PL, and AK) were allocated a selection of manuscript pairs and assessed the methods sections of the “before” and “after” versions. Each manuscript pair was assessed by three data extractors CS assessed the titles of all 20 manuscripts.

The assessors checked whether the “before” version submitted to Penelope.ai contained adequate information for each item in the methods section of the appropriate reporting checklist. Each item was assessed as present, absent, unclear/partial, or not applicable to that manuscript.

The assessors then checked whether the “after” version submitted to *BMJ Open* contained the same information as the “before” version or whether the author had added information. Each item was assessed as no change or added information.

CS collated and harmonised the data from the five assessors. In the “before” manuscripts, each item was given an overall code. Items were coded “1” if the item was partially reported or missing and “0” if there was adequate information:

- If two of the three assessments agreed, the majority assessment dictated the overall code.
- If the assessors disagreed and one or two assessors selected “not applicable,” the item was coded “0” indicating there was adequate information for that item.
- If the assessors each gave a different assessment and none selected “not applicable,” the item was coded “1” to indicate information was missing.

In the “after” manuscripts, each item given a rating of “1” (information missing) in the “before” version was given an overall code of “1” if information had been added or “0” if there was no change according to the majority assessment.

For each manuscript pair, we counted the number of items reported adequately in the “before” version and the number of items with more information in the “after” version. As each reporting guideline has a different number of items, we report these counts as percentages.

## **Ethics and consent**

In accordance with the University of Oxford's policy on the ethical conduct of research involving human participants and personal data (41), ethical approval and informed consent were not required. We used data collected as part of our partner PNL P Ltd.'s optional manuscript checking service, and during the normal course of *BMJ Open*'s editorial procedures. In accordance with the personal data protection policies of our partners, all data was anonymised before it was shared with the research team.

## Results

### 1. Individual user feedback

Between 16 January 2018 and 6 November 2019, 10,729 Penelope.ai users were recommended a reporting checklist on GoodReports.org as part of their feedback. Nearly 40% (4,182/10,729) of these users clicked the link to visit the GoodReports.org. All 10,729 users were sent an email survey one day after using Penelope.ai asking about their experience of using GoodReports.org. We received 623 responses.

#### Usefulness ratings

Most of the responders (599/623) answered the question "How useful did you find the checklist?" Figure 3 shows the distribution of ratings.

Only 176/599 of responders who answered question 1 (30%) rated GoodReports.org 7 or below for usefulness. Table 2 shows the responses of the 159/176 responders who explained why they gave this lower rating using our multiple-choice options and free text.

Table 2: Users' reasons for rating [www.goodreports.org](http://www.goodreports.org) 7/10 or lower. Users could select more than one multiple-choice option.

Why did you rate <a href="http://www.goodreports.org">www.goodreports.org</a> a [e.g., 6]?	
The checklist was too long	62/159 (39%)
The checklist was confusing	51/159 (32%)
The checklist items were not relevant to my work	50/159 (31%)
The website was confusing	6/159 (4%)
Other Free text answers	20/159 (13%)
<p><i>too difficult to complete the list points</i></p> <p><i>we have already done the Prisma checklist. These two checklists overlap each other.</i></p> <p><i>no clue what you're talking about. Nobody asked me to do a checklist.</i></p> <p><i>some of the checklist items were confusing</i></p> <p><i>some items are not relevant</i></p> <p><i>already addressed in my paper.</i></p> <p><i>checklist assume too much about the nature of 'good' work</i></p> <p><i>check list is the same as the one on the journal guidelines</i></p> <p><i>completed the whole form and then when I clicked the button at the end it deleted all my answers</i></p> <p><i>some of the checklist were not relevant to my work and i thought it would have looked at my discussion detail</i></p> <p><i>the checklist is still rather broad</i></p> <p><i>the checklist mentioned several items which were included in the article (e.g. corresponding author, subheadings)</i></p> <p><i>checklist mentioned that items were missing when they were present but with a slightly different spelling e.g. Conflicts of Interest instead of Conflicts of Interests</i></p> <p><i>not listed in the journal's Instructions to authors</i></p> <p><i>was incorrect</i></p> <p><i>good</i></p> <p><i>some items were not available to my article.</i></p>	

*s a decent rating*  
*ir, the checklist differs with different countries*  
*REQ checklist was used.*

## Suggestions for improvement

274/623 respondents (44%) responded to the question "How could we make [www.goodreports.org](http://www.goodreports.org) more useful?" Of these, 71 (26%) gave a neutral response with no suggestion (e.g., "Not sure"), 57 (21%) were general compliments (e.g., "Easy to navigate and very useful"), 50 (18%) were comments about the workflow through which they had encountered GoodReports.org (e.g., from *BMJ Open* or via the Penelope.ai manuscript checker), and 6 (2%) were criticisms of reporting checklists in general (e.g., "Blind checklists are not relevant to most work"). We did not consider these responses further.

The remaining 90/274 (33%) responses included 93 actionable suggestions for improvement, with some responses including more than one suggestion. Table 3 shows the broad themes in these suggestions and representative comments.

Table 3. Themes and representative free-text responses (in italics) to the question "How could we make [www.goodreports.org](http://www.goodreports.org) more useful?"

Themes	Number of responses
<p>Simplification (e.g., shorten the checklists and their items and clarify the item's wording)</p> <p><i>By simplifying it and much shortening</i></p> <p><i>A bit shorter would be good. Too many questions on data monitoring and storage</i></p> <p><i>The resulting interface is not clear enough</i></p>	<p>26/93 (28%)</p>
<p>Applicability (e.g., include a wider variety of guidelines and remove non-applicable reporting items for that study)</p> <p><i>can include more checklists that are commonly used.</i></p> <p><i>provide more specific checklist items and give example if it is possible, to avoid us to get confused maybe by providing feedback</i></p> <p><i>More options for a specific report. Our report was an animal based cadaveric study looking at accuracy of drill guides. We were unsure which category it should fall under.</i></p> <p><i>There should be an explanation that not all questions might be applicable.</i></p>	<p>17/93 (18%)</p>
<p>Give more information (e.g., examples of good reporting, definitions of reporting items, and more instruction on how to use the checklist)</p> <p><i>Give more detailed information about filling out checklist items. Some items are not very clear.</i></p> <p><i>Definitions of all of the terms as hyperlinks.</i></p> <p><i>Add boxes on the checklist to add the corresponding page number</i></p> <p><i>have a key to explain basic terms such as sensitivity analysis for first time users, forest plot</i></p> <p><i>Online video tutorials or workshops would help a lot</i></p>	<p>15/93 (16%)</p>
<p>Technical improvements to the GoodReports website</p> <p><i>a window that allows you to view your manuscript in real time while going through the checklist will be highly valuable.</i></p> <p><i>Put in the checklist automatically</i></p>	<p>11/93 (12%)</p>

<i>Mark in the text what corresponds to each item in the list</i>	
<p>Promotion (e.g., suggestions to reach more authors and at earlier stage of writing)</p> <p><i>Easily accessible... Inform through social media platforms</i></p> <p><i>make available for all journal submissions</i></p> <p><i>You could make more visible to the junior researchers.</i></p> <p><i>Dissemination in universities and research centers</i></p> <p><i>Encourage people to use the criteria early in the writing process (I have, which probably is why I only changed one thing)</i></p>	<p>10/93 (11%)</p>
<p>Suggestions for specific guidelines</p> <p><i>For a systematic review, the PRISMA guidelines can also be used rather than the MOOSE.</i></p>	<p>2/93 (2%)</p>
<p>Unclear (we were not certain what the user meant)</p> <p><i>by making/and more highlighting the specific parts for readers.</i></p> <p><i>integrating them in reporting guidelines</i></p> <p><i>Reducing the burden of comparison a manuscript with others</i></p>	<p>12/93 (13%)</p>

### User feedback on completeness

Between 4 December 2018 and 6 November 2019, we received 267 responses to the question “After using the checklist did you make any changes to your manuscript?” Most respondents (198/267, 74%) said they had made changes, and 69 (26%) said they had not. Table 4 summarises the reasons given for not making changes and includes examples of user responses about changes they had made.

Table 4: Themes and representative comments (in italic) about changes made or reasons for not making changes.

Themes	Number of responses
<b>Respondents that <i>had</i> made changes</b>	<b>198/267 (74%)</b>
<p>Provided reasonable detail about what they had changed</p> <ul style="list-style-type: none"> <li>• <i>Some precision on the definition of the dimension of of intervention we were studying</i></li> <li>• <i>extra detail added regarding blinding/allocation of intervention</i></li> <li>• <i>I added the consent form, protocol contributor information, role of study sponsor, adherence efforts, criteria for discontinuation, and updated safety measures and data security descriptions.</i></li> <li>• <i>Added more detail about the conceptual framework used</i></li> <li>• <i>Add data management and confidentiality and harms to the manuscript</i></li> <li>• <i>included how the sample size was determined</i></li> <li>• <i>mainly just adding more detail e.g. about method of randomisation</i></li> </ul>	51/198 (26%)
<p>Provided non-specific information about what they had changed</p> <ul style="list-style-type: none"> <li>• <i>The structure of manuscript and details in each section</i></li> <li>• <i>Some of the items that are stated in the STROBE but were not written in the original manuscript</i></li> <li>• <i>Made sure that each sub-title conformed to the checklist suggested nomenclature</i></li> <li>• <i>A lot! Added info and checked other components</i></li> <li>• <i>More details that were missing in some sections</i></li> </ul>	115/198 (58%)
<p>Did not provide details about what they had changed</p>	32/198 (16%)
<b>Respondents that <i>had not</i> made changes</b>	<b>69/267 (26%)</b>

Had already followed a reporting guideline when writing	16/69 (23%)
No edits required because they had already reported all items (but did not explicitly mention following a reporting guideline)	31/69 (45%)
Did not want to <ul style="list-style-type: none"> <li>• <i>I did not have the time finish it so I changed a journal</i></li> <li>• <i>too much TROUBLE</i></li> <li>• <i>I don't accept the changes</i></li> </ul>	7/66 (10%)
Unclear (e.g., "It's OK")	7/66 (10%)
No answer	8/66 (12%)

## 2. Reliability of the questionnaire in helping authors find the most appropriate reporting guideline for their work

Between 25 January 2018 and 16 February 2019, 5,831 authors submitting to *BMJ Open* elected to check their work with Penelope.ai before completing submission. We randomly selected 100 of these manuscripts and compared our recommendations of reporting guidelines with that recommended by the checklist finder questionnaire.

We agreed with 73/100 of the questionnaire recommendations: 57/100 manuscripts were recommended an appropriate guideline and 16/100 were correctly told that no appropriate guideline existed.

We disagreed with 27 of the questionnaire recommendations: 21 cases where we thought the wrong checklist had been recommended and 6 cases where we thought there was an appropriate checklist, but none was recommended

In the 27 cases where we assessed that the wrong recommendation had been made, we compared the authors' responses to the questionnaire with the associated manuscript for clues as to how the questionnaire could be improved. These results are summarized in Table 5.

Table 5. Summary of possible reasons for inappropriate recommendations, with potential solutions

Reporting guideline inappropriately recommended	# of users	Possible reason for mismatch	Potential solution
SRQR for reporting qualitative research	7	All responded that their study used exclusively qualitative data. However, although most of the studies did collect qualitative data, they also collected quantitative data and used a standard design that would have better fit another reporting guideline (e.g., CONSORT or STROBE). As the current question in the finder is very long, authors may have missed the word “exclusively.”	Split the question into two to accommodate studies with both qualitative and quantitative data. Allow the finder to recommend a combination of reporting guidelines.
SPIRIT for reporting protocols of randomised trials	3	All responded that they were reporting a protocol of a clinical trial, but their protocol was for a different study design. There are only reporting guidelines for protocols for two study designs (clinical trials and systematic reviews). Authors writing other kinds of protocol may have decided to choose one of the presented designs rather than selecting “other” on the final question.	Give more useful advice to authors writing protocols for which no guideline currently exists or is openly available. It may be more useful to direct authors to the design’s reporting guideline instead, with an instruction to focus on the methods.
STARD for reporting diagnostic test accuracy studies	3	All responded that they were investigating a diagnostic test.  Two were studies that focused on tests and screening but were not primary studies of	Make the question directing people to STARD more precise, to be clear that it does not refer to all studies focused on tests and screening.

		<p>diagnostic test accuracy (DTA). They would have been better matched by STROBE.</p> <p>The third was a meta-analysis of DTA studies. They were not matched with PRISMA as they had not used a systematic review design.</p>	<p>Include a question to identify studies that use meta-analysis without systematic review.</p>
STROBE cohort	2	<p>Both studies were correctly identified as using observational study designs, with mismatch in one case on which of the STROBE sub-designs applied.</p> <p>One study was a cross-sectional study of a previous research cohort.</p> <p>The second was a retrospective cohort study on the results of a screening test which we incorrectly judged should have been matched with STARD.</p>	<p>Include clearer explanations, potentially with graphics, to help authors identify which type of observational study they have done.</p>
CONSORT for reporting randomised trials	2	<p>The authors of the first study indicated that they used an experimental design. However, the study was a survey with no investigated intervention.</p> <p>The authors of the second study correctly identified their work as experimental. However, they responded that two groups were being compared when they used a single-arm study design.</p>	<p>Include a clearer explanation of the difference between experimental and observational research.</p> <p>Direct authors of phase 1 trials better.</p>
CHEERS for economic	1	<p>Although this was a cohort study the study participants were selected on basis of resource</p>	<p>Clarify that the question regarding economic aspects</p>

evaluations		use. The authors therefore indicated that the study was looking at economic aspects of healthcare.	only refers to primary outcomes.
MOOSE for meta-analysis of observational studies	1	Although the author's responses were reasonable, the study included meta-analyses of both trials and observational research. PRISMA would have been a better fit.	Include advice on what to do if your manuscript reports multiple study types.
PRISMA-P for protocols of systematic reviews	1	Although the authors responded that they were writing a protocol for a systematic review, they were planning a "realist" review. There is no appropriate reporting guideline for this design.	Give more useful advice to authors writing protocols for which no guideline currently exists or is openly available.
STROBE case-control	1	The study was correctly identified as observational research, with mismatch on which of the 3 STROBE sub-designs applied. This before-and-after study used a cohort design, not a case-control design as selected. As the study related to a new care management system, SQUIRE might also have been helpful.	Include clearer explanations, potentially with graphics, to help authors identify which type of observational study they have done.  Include advice on what to do if more than one reporting guideline would apply to your manuscript.
<b>Recommendation that there is no appropriate checklist</b>	<b># of users</b>		
<b>Our recommendation</b>			
SPIRIT	2	1 of the manuscripts was a pilot randomised trial,	Expand explanation of what a

		<p>but the author response indicated that their article was a protocol.</p> <p>The authors of the other manuscript identified it as a protocol, but then selected “other” instead of clinical trial. The trial was community-based not clinic-based, so their response was technically correct</p>	<p>protocol is.</p> <p>Change “clinical trial” to “randomised trial” in the questionnaire</p>
CONSORT	1	<p>The response did not logically match the study design. The authors responded to the initial screening question that they were not doing health or biomedical research, although the article reported secondary analyses of a randomised clinical trial</p>	
STROBE Cohort	1	<p>The authors chose "other" from options of article type (rather than case report, protocol, systematic review, or research article). However, the work presented initial findings of a birth cohort</p>	<p>Clarify what ‘research’ means, perhaps by adding examples (e.g. experimental or observational).</p> <p>Add more helpful feedback than "no guideline" for these responses.</p>
STROBE Cross-sectional	1	<p>The authors did not give any further response after saying that the study collected data from people.</p> <p>They may have struggled with the next question, "What are you investigating?" As they conducted observational/epidemiological research, the correct option was “factors affecting health.”</p> <p>This option may be too vague.</p>	<p>Clarify what ‘factors affecting health’ refers to, to make the connection to epidemiological/observational research clearer.</p>

STROBE Case-control	1	The authors correctly responded they were writing a protocol and provided no further responses. Their paper was a protocol for a case-control study, so STROBE case-control would have been a good fit	Give more useful advice to authors writing protocols for which no guideline currently exists or is openly available
---------------------	---	--	---

Figure 4 shows how we selected manuscripts to observe submission quality when GoodReports is used, focusing on whether a checklist was submitted and completeness of reporting. Over our six data collection days in May 2018, 217 newly submitted manuscripts were checked by the *BMJ Open* editorial team. Five manuscripts were excluded because the author withdrew their submission between when the manuscript was checked and when *BMJ Open* exported the data to share with us. As *BMJ Open* deletes submission data at withdrawal, we could not check whether these manuscripts had used Penelope.ai. We also excluded two manuscripts that were flagged as duplicate submissions that had already been checked once before by *BMJ Open* staff.

We matched 74 of the remaining 210 submissions (35%) to Penelope.ai uploads. We excluded manuscripts from 5 authors who did not view their feedback (which may have included a reporting guideline recommendation) on Penelope.ai. We therefore analysed compliance with checklist submission using 205 manuscripts, 69 that were exposed to GoodReports via Penelope.ai and 136 that were not.

The editorial office checked each submission to determine whether a checklist was required and, if so, whether it had been submitted. Of the authors that did not use Penelope.ai, *BMJ Open's* editorial team had to chase 41/136 (30%) of them for a checklist. Of the authors that did use Penelope.ai, the editorial team had to chase 10/69 (14%).

#### 4. Completeness of reporting before and after using a GoodReports reporting checklist

Of the 69 manuscripts that had used Penelope.ai, 10 were excluded because the GoodReports questionnaire did not make a checklist recommendation, 16 because they did not submit a checklist to the journal, 22 because the checklist submitted alongside the manuscript had not come from GoodReports.org, and 1 because it was submitted with the wrong GoodReports checklist. We therefore assessed 20 manuscripts (Figure 4). Table 6 compares the completeness of reporting in the manuscript version submitted to Penelope.ai before completing a GoodReports reporting checklist with the version submitted to *BMJ Open* after completing the recommended checklist

Table 6: Completeness of reporting in manuscripts before and after completing a GoodReports checklist

	Applicable reporting checklist	Number of method items in checklist	Items reported <i>before</i> completing a reporting checklist	Items reported (or reported more fully) <i>after</i> completing a reporting checklist	Number of reporting items improved
1	PRISMA-P	15	12	12	0
2	PRISMA-P	15	12	14	2
3	PRISMA-P	15	13	13	0
4	SRQR	11	6	7	1
5	SPIRIT	25	18	18	0
6	SPIRIT	25	12	12	0
7	SPIRIT	25	14	20	6
8	SPIRIT	25	15	16	1
9	CONSORT	16	10	11	1
10	STROBE Cohort	14	4	4	0
11	STROBE Cohort	14	10	10	0
12	STROBE Cohort	14	10	10	0
13	STROBE Cross-sectional	13	10	10	0
14	STROBE Cross-sectional	13	5	5	0
15	STROBE Cross-sectional	13	6	6	0
16	TRIPOD	18	11	11	0
17	STARD	17	8	8	0

18	MOOSE	8	6	6	0
19	MOOSE	8	2	2	0
20	SQUIRE	8	0	0	0

CONSORT: Consolidated Standards of Reporting Trials; MOOSE: Meta-analyses Of Observational Studies in Epidemiology; PRISMA-P: Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols; SRQR: Standards for Reporting Qualitative Research; SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials; STARD: Standards for Reporting Diagnostic Accuracy; STROBE: STrengthening the Reporting of OBservational studies in Epidemiology (sub-checklists for cohort, cross-sectional, and case-control studies); TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

Five of the 20 (25%) manuscripts improved their methodological reporting after having completed a reporting checklist on GoodReports.org. Of these, 3 added information for 1 reporting item, 1 added information for 2 items, and 1 added information for 6 items. Three of the 5 (60%) improved manuscripts were protocols, whereas 7/20 (35%) of the full sample were protocols.

No manuscript completely described all methodological reporting items from the recommended reporting checklist, either before or after completing a checklist. One manuscript failed to address any items from its checklist. On average, manuscripts described 57% (SD 22%) of necessary reporting items before completing the checklist and 60% (SD 23%) after the checklist.

Eight manuscripts in the before group (40%, n=20) had titles that met guideline requirements, 6 (30%) that partially met requirements, and 6 (30%) that contained no information stipulated in the guideline item for title. Most guidelines stipulate describing the study design in the title as the main or only requirement. Two manuscripts in the “after” group improved their titles by adding information.

## Discussion

The results encourage us to think that using a questionnaire to direct authors to an appropriate reporting guideline available in a standardised format on one website may be an effective way to increase their use.

### 1. Individual user feedback

The results from our user survey suggested that users found the GoodReports website usable and useful. Most respondents (74%) reported making edits to their manuscript after using it and could back this up with examples of what they changed. The common objections and suggestions paint a coherent picture:

users would find checklists more useful if they were shorter, easier to understand, and more applicable. Users also suggested that GoodReports be implemented at an earlier stage of writing.

## **2. Questionnaire reliability**

Our questionnaire gave appropriate advice to 73% of users. This is an improvement on the previous version of the questionnaire (16). A previous study of authors submitting to four BioMed Central journals found that the correct version of a required checklist was submitted by 31% of authors (70/224) who were given no further assistance and 40% of authors (86/217) aided by a previous version of the reporting guideline finder (30).

The questionnaire can be improved further. Sometimes the questionnaire had no recommendation where we thought there was an appropriate checklist available. Finding the right reporting guideline is not a straightforward task for either authors or the journals that endorse or even require authors to submit them. For example, a study of the effectiveness of a web-based tool to improve the reporting of randomised trials revealed that editorial staff were often unable to correctly identify a randomised trial based on what was reported in submitted manuscripts (31).

A common issue was authors responding that they were collecting exclusively qualitative data when they were either collecting exclusively quantitative data or both qualitative and quantitative data. As the question is quite long, people may have missed or misunderstood the word “exclusively.” We could split the question about qualitative data into two, first asking whether they collected qualitative data and then whether they collected quantitative data. Those answering yes to both questions could be redirected onto the main questionnaire journey and provided a link to the reporting guideline for reporting qualitative research in addition to a main guideline recommendation.

Another common issue was authors of protocols either receiving no recommendation or an inappropriate recommendation. There are few reporting guidelines available for writing protocols for different study designs, and those that do exist often have restrictive usage licences. As a workaround, we will adjust the questionnaire to recommend that authors of protocols without a specific reporting guideline could use the appropriate reporting guideline for a completed study of the same design to guide them when writing their manuscripts. We will also provide links to the protocol reporting guideline publication if there is one.

As we plan to add around 30 more reporting guidelines to the GoodReports database, future versions of the questionnaire will offer authors lists of options to help identify guidelines based on the design (e.g., type of trial or observational study), type of intervention or exposure (e.g., nutrition or psychological intervention), type of outcomes measured (e.g., economic or health equity), and focus of desired healthcare improvement (e.g., health policy or service delivery). Common study designs that cannot be matched to a reporting guideline could help direct future guideline development by indicating the largest need.

## **3. Rates of submission**

Partnering with *BMJ Open* and Penelope.ai was an effective way to attract users to [www.goodreports.org](http://www.goodreports.org) and observe author behaviour. Authors who chose to use Penelope.ai for a paper check were more likely to submit a recommended reporting checklist with their submission, giving the editorial team one less thing to chase. A before-and-after study across four speciality medical research journals testing the previous version of the GoodReports questionnaire also found that its use during submission was associated with improved author identification of the relevant reporting guidelines for their study type (17). We cannot comment on causality here, as the data are all observational.

We found a discrepancy between how many manuscripts were identified as missing checklist submissions by the *BMJ Open* staff (10/69) and our team (16/69), when focusing on manuscripts that had been checked by Penelope.ai. We were unable to check the *BMJ Open* staff decisions for the rest of the manuscripts submitted in the 6-day window. The *BMJ Open* staff were unaware of whether an author had used Penelope.ai. It is therefore likely that they misclassified checklist needs for a similar proportion of manuscripts that had not been submitted to Penelope.ai. This is useful data for journals that plan to ask their staff to enforce reporting guideline policies. Journal staff may benefit from tailored training in identifying a manuscript's study design and an appropriate reporting guideline. This also reminds us that it is insufficient to rely on one group within the research community to enforce reporting standards.

#### **4. Completeness of reporting**

A quarter of authors made changes to their title or methods section after completing a reporting checklist. Examples of significant changes included the addition of a paragraph on data management and a paragraph on power calculation. However, any change at all was rare, and the maximum changes recorded was six in one of the twenty manuscripts. On average, 40% of items were still missing from the methods section after completing an appropriate reporting checklist. The three systematic review protocols were the best reported manuscripts in our sample, but still missed several items after completing the PRISMA-P checklist. One quality improvement study had completed the SQUIRE checklist, but our assessors were unable to find any of the required items in the methods section. It is possible that the items had been reported in another section of the manuscript, which the assessors did not see. These results contrast with the user feedback study, where three-quarters of respondents said they had made changes after using a reporting checklist.

In our small sample, submission of a reporting checklist with an article did not indicate the article had been completely reported. The higher rate of checklist submission by authors who received a reporting guideline recommendation did not correspond with a higher rate of more complete manuscripts.

These results suggest to us that recommending reporting guidelines at the point of submission may be too late. By this stage, authors may not have the time, ability, or motivation to make substantial edits and clear changes with their co-authors. This problem will particularly apply to the methods section of completed studies. If key elements of the methods specified in the reporting checklist have not been carried out or recorded, they cannot be reported.

Reporting guidelines are designed to be used to direct early writing, not just as checklists after writing is finished. Authors may also be more likely to react positively to reporting guidance even earlier in the research process, at funding application, protocol stage, or before data collection, rather than when writing up results. Seeing the essential information needed in a good study report early in the research process means researchers can adjust their plans if needed.

## **Limitations**

The data collected were all observational, and sample sizes were determined by journal editorial systems and our team's availability to assess manuscripts. We cannot state whether GoodReports is an effective intervention until we test it in a randomised trial (32).

The author feedback data, including ideas on how to improve the tool, was collected from authors who chose to use Penelope.ai and try out the GoodReports tool. These people would be broadly representative of *BMJ Open* and Penelope.ai's other journal clients, but not representative of all authors. However, the respondents were likely to be more representative than previous author experience surveys, which have generally been small, focused on a single reporting guideline, and included mainly European and US respondents (33-36).

For pragmatic reasons, we restricted our assessment of reporting completeness to the methods section and title, which may have skewed the effect of exposure to GoodReports. However, the primary purpose of the manuscript completeness assessment study was to pilot the assessment task with the team. This will inform the assessment protocol for a future randomised trial where we will assess both the methods and results sections (32).

The assessors could not be blinded because the "before" file format was different from the "after" file format. Any potential for bias was mitigated by the objective outcome measure – there was no ambiguity over whether the two versions of text were identical or whether text had been added.

## **Next development steps**

Following user feedback asking for more clarification and help in applying reporting guidelines, we will try to focus GoodReports users' attention on important, non-generic reporting items, such as randomisation procedures. Many guidelines include generic recommendations, such as what to include in the introduction and discussion. This makes guidelines very long and may give users the false impression that the checklist is too general. We will also make existing reporting items easier to understand by adding examples of good reporting and definitions of key terms. We will investigate how to provide tailored guidance that combines items from more than one reporting guideline.

We will continue to improve the questionnaire and will add more standalone guidelines and more extensions of existing guidelines to the database. We will test different ways of wording and implementing the questionnaire to ensure that more users reach an appropriate recommendation.

Instead of capturing authors at the point of submission to a journal, we will aim to reach them earlier in the research process, at least before they have started to write up their study. Potential avenues include contacting authors that publish Registered Reports (37), preprints (38), and protocols; advertising in author education programmes; and working with researcher support organisations such as Author Aid (39). To better support authors beginning to write, we will investigate delivering reporting guidance in article templates.

We intend to use interviews, surveys, and focus groups to inform recruitment techniques for the randomised trial. We will test the acceptability of both the GoodReports and control interventions to ensure participants will be happy to be randomised.

We were heartened by one user's comment: "I will refer to all checklists to write my report clearly and fully next time." We are optimistic that the tools we are developing will help authors, increase the adoption of reporting guidelines, and, ultimately, reduce research waste.

## List Of Abbreviations

EQUATOR: Enhancing Quality and Transparency of Health Research

PRISMA-P: Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols

SRQR: Standards for Reporting Qualitative Research

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

CONSORT: Consolidated Standards of Reporting Trials

STROBE: STrengthening the Reporting of OBservational studies in Epidemiology

TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

STARD: Standards for Reporting Diagnostic Accuracy

MOOSE: Meta-analyses Of Observational Studies in Epidemiology

## Declarations

### Ethics approval and consent to participate

In accordance with the University of Oxford's policy on the ethical conduct of research involving human participants and personal data (41), ethical approval and informed consent were not required. We used data collected as part of our partner PNLP Ltd.'s optional manuscript checking service, and during the normal course of BMJ Open's editorial procedures. In accordance with the personal data protection policies of our partners, all data was anonymised before it was shared with the research team.

Before using Penelope.ai, users were asked for consent for their manuscript and feedback to be used for research and evaluation purposes. Penelope.ai had Data Sharing agreements in place with the University of Oxford and *BMJ Open* to allow the sharing of unpublished, de-identified excerpts of manuscripts during the assessment period.'

### **Consent for publication**

Not applicable

### **Availability of data and materials**

The datasets generated and/or analysed during the study are available on the Open Science Framework GoodReports development site: <https://osf.io/2pejk/>

1. Individual user feedback: <https://osf.io/h35aw/>
2. Reliability of the questionnaire in helping authors find the most appropriate reporting guideline for their work: <https://osf.io/fmq9s/>
3. Exposure to GoodReports.org and rates of submission of a completed reporting checklist: <https://osf.io/6ajh3/>
4. Exposure to GoodReports.org and completeness of reporting (before-and-after study): <https://osf.io/u4fn7/>

Protocol for assessors: <https://osf.io/t8jpm/>

Sample data collection form: <https://osf.io/jnk8h/>

### **Competing interests**

CS, JdB, JH, PL, PD, and MS are members of the UK EQUATOR Centre, which is part of the EQUATOR Network. The EQUATOR Network's goal is to improve the quality and transparency of health research, primarily through the promotion of reporting guidelines.

JH is founder and director of PNL P Ltd, the company that makes Penelope.ai.

### **Funding**

This publication presents independent research funded by the National Institute for Health Research (NIHR), Medical Research Council, and Cancer Research UK. The views expressed are those of the author(s) and not necessarily those of Cancer Research UK, the Medical Research Council, the NHS, the NIHR, or the Department of Health and Social Care.

CS is funded by the Medical Research Council: when conducting the study: MR/J00488X/1; when writing up the study: MR/S036741/1

JH is funded by the Medical Research Council MR/J00488X/1 and Cancer Research UK (C49297/A29084)

JdB is funded by Cancer Research UK (C5529/A16895 and C49297/A27294)

PD is supported by the NIHR Biomedical Research Centre, Oxford.

PL is supported by the NIHR Biomedical Research Centre, Oxford, and Cancer Research UK (C49297/A27294)

MS is funded by Cancer Research UK (C49297/A27294).

### **Author contribution statement**

Conceptualization: JH, CS, Data curation: JH, CS, Formal analysis: JH, CS, Investigation, Methodology: JH, CS, JdB, PD, PL, MS; Project administration: CS; Software: JH; Supervision: CS, Validation: JH; Writing - original draft: CS, JH, Writing - review & editing: CS, JH, JdB. PD, PL, MS.

### **Acknowledgements**

We wish to acknowledge the staff at *BMJ Open*, especially their Editor-in-Chief, Adrian Aldcroft, for providing data from article submissions.

Thanks to the members of the original GoodReports working group convened and led by Iveta Simera and Jason Roberts in 2015. Other members of the group who contributed significantly to the development of GoodReports were Larissa Shamseer, David Moher, and Daniel Shanahan.

To Doug Altman, Gary Collins, and members of the UK EQUATOR Centre team, many thanks for providing feedback and encouragement during the development of the first and subsequent versions of the checklist finder and during development of the final GoodReports prototype.

Thanks to Anna Koroleva for her contributions to the protocol and data collection for the before and after study.

Additional thanks to Susan Galandiuk, Editor-in-Chief *Diseases of the Colon and Rectum*, and Adrian Aldcroft, mentioned above, for their letters in support of our grant application to develop and test GoodReports in a randomised trial, and for committing to the involvement of their journals in future work.

## **References**

1. Jin Y, Sanger N, Shams I, Luo C, Shahid H, Li G, et al. Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? – A systematic review of reviews: an update. *J Multidiscip Healthc*. 2018 Sep 27; 11:495–510.
2. The EQUATOR Network | Enhancing the QUALity and Transparency Of Health Research [Internet]. [cited 2020 Feb 14]. Available from: <https://www.equator->

network.org/ 3. National Centre for the Replacement, Refinement and Reduction of Animals in Research [Internet] [cited 2020 Nov 09]. Available from: <https://www.nc3rs.org.uk/> 4. Cochrane [Internet] [cited 2020 Nov 09]. Available from: <https://www.cochrane.org/> 5. International Committee of Medical Journal Editors [Internet] [cited 2020 Nov 09]. Available from: <http://www.icmje.org/> 6. World Association of Medical Editors [Internet] [cited 2020 Nov 09]. Available from: <https://www.wame.org/> 7. European Association of Science Editors [Internet] [cited 2020 Nov 09]. Available from: <https://ease.org.uk/> 8. Committee for Publication Ethics [Internet] [cited 2020 Nov 09]. Available from: <https://publicationethics.org/> 9. Botos J. Reported use of reporting guidelines among JNCI: Journal of the National Cancer Institute authors, editorial outcomes, and reviewer ratings related to adherence to guidelines and clarity of presentation. *Res Integr Peer Rev*. 2018; 3:7. 10. Hair, K., Macleod, M.R., Sena, E. et al. A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines (IICARus). *Res Integr Peer Rev* 4, 12 (2019). <https://doi.org/10.1186/s41073-019-0069-34>. 11. Hopewell, S., Boutron, I., Altman, D.G. et al. Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: results of a randomised controlled trial. *BMC Med* 14, 199 (2016). <https://doi.org/10.1186/s12916-016-0736-x> 12. Blanco D, Altman D, Moher D, Boutron I, Kirkham JJ, Cobo E. Scoping review on interventions to improve adherence to reporting guidelines in health research. *BMJ Open*. 2019 May;9(5): e026589. 13. GoodReports.org [Internet]. [cited 2020 Feb 14]. Available from: <https://www.goodreports.org> 14. Penelope.ai [Internet]. Penelope.ai. [cited 2020 Feb 14]. Available from: <https://www.penelope.ai> 15. Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, Browne WJ, Clark A, Cuthill IC, Dirnagl U, Emerson M, Garner P, Holgate ST, Howells DW, Karp NA, Lazic SE, Lidster K, MacCallum CJ, Macleod M, Pearl EJ, Petersen O, Rawle F, Peynolds P, Rooney K, Sena ES, Silberberg SD, Steckler T and Wurbel H. The ARRIVE Guidelines 2.0: updated guidelines for reporting animal research. 16. Gagnier JJ, Kienle G, Altman DG, et al The CARE guidelines: consensus-based clinical case reporting guideline development *Case Reports* 2013;2013:bcr2013201554 17. Husereau Don, Drummond Michael, Petrou Stavros, Carswell Chris, Moher David, Greenberg Dan et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement *BMJ* 2013; 346 :f1049 18. Schulz KF, Altman DG, Moher D, for the CONSORT Group (2010) CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials. *PLoS Med* 7(3): e1000251. <https://doi.org/10.1371/journal.pmed.1000251> 19. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group*. *JAMA*. 2000; 283(15):2008-2012. 20. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7): e1000097. <https://doi.org/10.1371/journal.pmed.1000097> 21. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):1. 22. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. *Ann Intern Med*. 2013;158(3):200-207. 23. Ogrinc G, Davies L, Goodman D, et al SQUIRE 2.0

(Standards for Quality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process *BMJ Quality & Safety* 2016;25:986-992. 24. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med.* 2014;89(9):1245-1251. 25. Bossuyt Patrick M, Reitsma Johannes B, Bruns David E, Gatsonis Constantine A, Glasziou Paul P, Irwig Les et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies *BMJ* 2015; 351 :h5527 26. Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, et al. (2009) STrengthening the REporting of Genetic Association Studies (STREGA)— An Extension of the STROBE Statement. *PLoS Med* 6(2): e1000022. <https://doi.org/10.1371/journal.pmed.1000022> 27. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *PLoS Med* 4(10): e296. <https://doi.org/10.1371/journal.pmed.0040296> 28. Collins, G.S., Reitsma, J.B., Altman, D.G. et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med* 13, 1 (2015). <https://doi.org/10.1186/s12916-014-0241-z> 29. EQUATOR Reporting Guideline Decision Tree <https://www.equator-network.org/wp-content/uploads/2013/11/20160226-RG-decision-tree-for-Wizard-CC-BY-26-February-2016.pdf> [cited 2021 Feb 04] 30. Shanahan DR, Lopes de Sousa, Marshall DM. Simple decision-tree tool to facilitate author identification of reporting guidelines during submission: a before-after study. *Research Integrity and Peer Review.* 2017 Dec 18; 2:20 31. Hopewell, S., Boutron, I., Altman, D.G. et al. Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: results of a randomised controlled trial. *BMC Med* 14, 199 (2016). <https://doi.org/10.1186/s12916-016-0736-x> 32. Does providing authors with customized article templates including items from reporting guidelines improve completeness of reporting? A randomized trial [Internet]. [cited 2021 Jan 25] Available from: <https://osf.io/d8u76> 33. Sharp MK, Glonti K, Hren D. Online survey about the STROBE statement highlighted diverging views about its content, purpose, and value. *J Clin Epidemiol.* 2020 Jul;123:100-106. doi: 10.1016/j.jclinepi.2020.03.025. Epub 2020 Apr 4. PMID: 32259582. 34. Fuller T, Pearson M, Peters J, Anderson R (2015) What Affects Authors' and Editors' Use of Reporting Guidelines? Findings from an Online Survey and Qualitative Interviews. *PLoS ONE* 10(4): e0121585. <https://doi.org/10.1371/journal.pone.0121585> 35. Davies L, Batalden P, Davidoff F, Stevens D, Ogrinc G. The SQUIRE Guidelines: an evaluation from the field, 5 years post release. *BMJ Qual Saf.* 2015;24(12):769-775. doi:10.1136/bmjqs-2015-004116 36. Burford BJ, Welch V, Waters E, Tugwell P, Moher D, O'Neill J, et al. (2013) Testing the PRISMA-Equity 2012 Reporting Guideline: the Perspectives of Systematic Review Authors. *PLoS ONE* 8(10): e75122. <https://doi.org/10.1371/journal.pone.0075122> 37. Registered Reports [Internet]. [cited 2020 Feb 14]. Available from: <https://cos.io/rr/> 38. The YODA Project | medRxiv [Internet]. [cited 2020 Feb 14]. Available from: <https://yoda.yale.edu/medrxiv> 39. AuthorAID - Home [Internet]. [cited 2020 Feb 14]. Available from: <https://www.authoraid.info/en/> 40. GoodReports feedback survey: [cited 2021 March 5] Available from: <https://pnlp.typeform.com/to/NM1epn#guideline=xxxxx&pid=xxxxx> 41. University of Oxford's Policy on the ethical conduct of research involving human participants and personal data [cited 2021 March 8] Available from: <https://researchsupport.admin.ox.ac.uk/governance/ethics/committees/policy>

# Figures

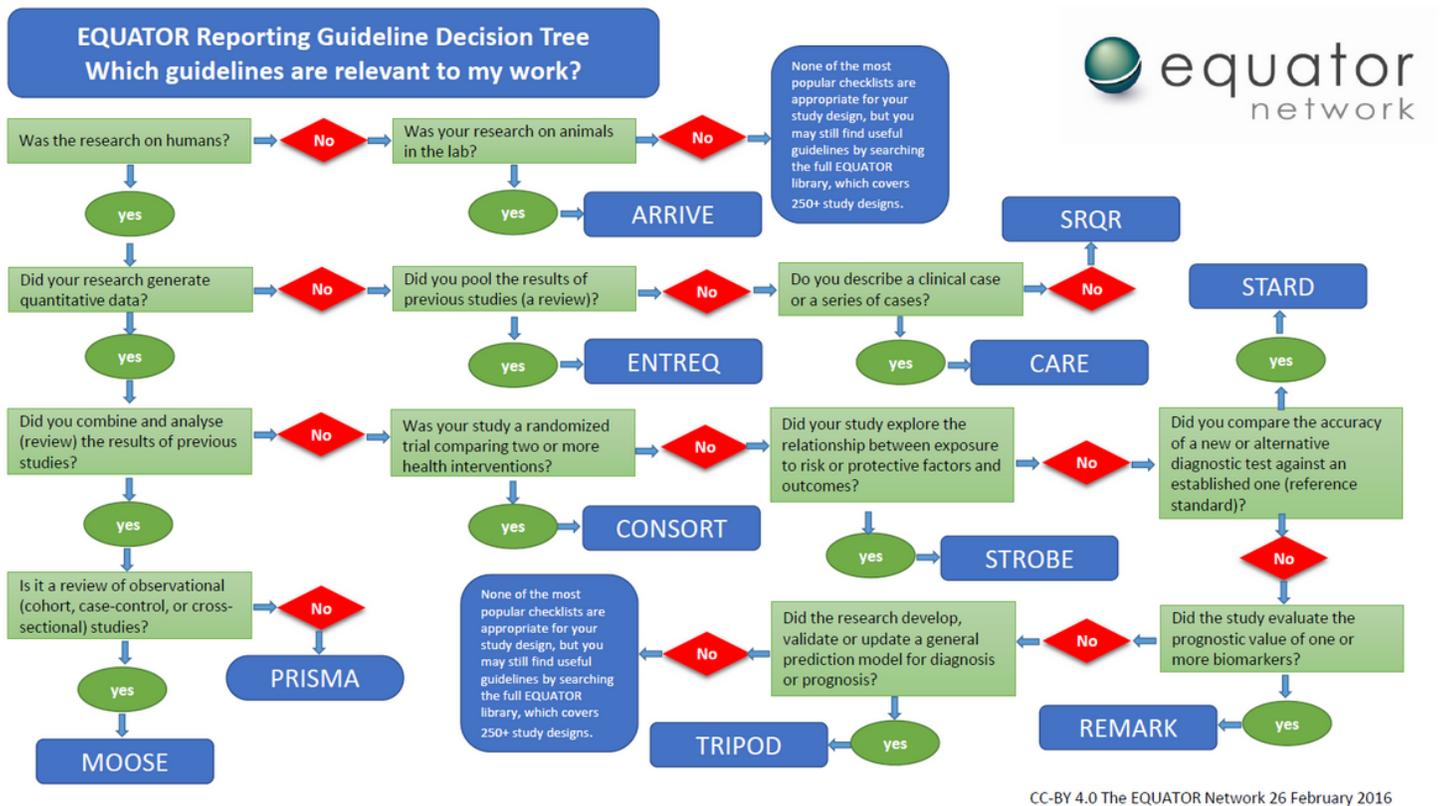
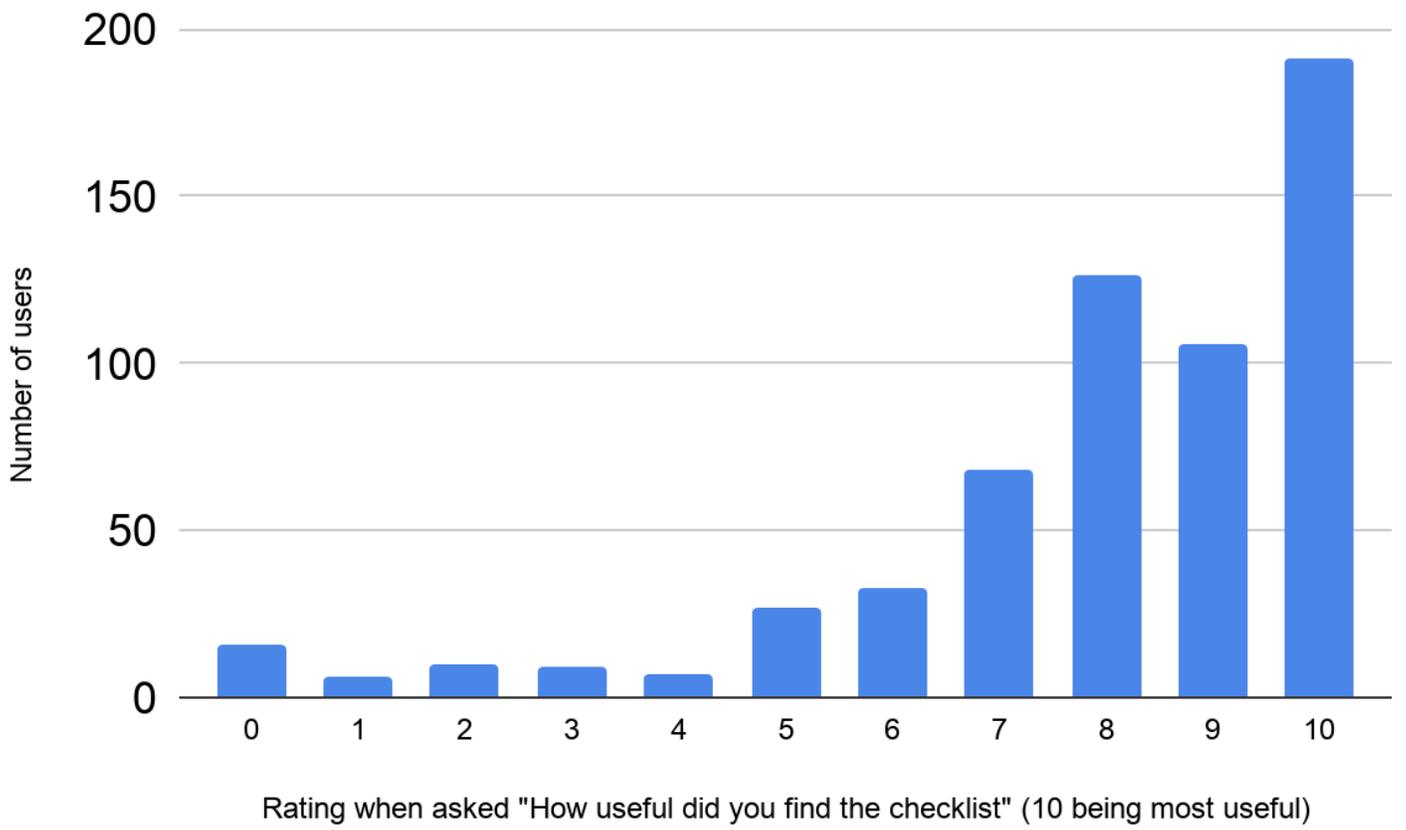


Figure 1

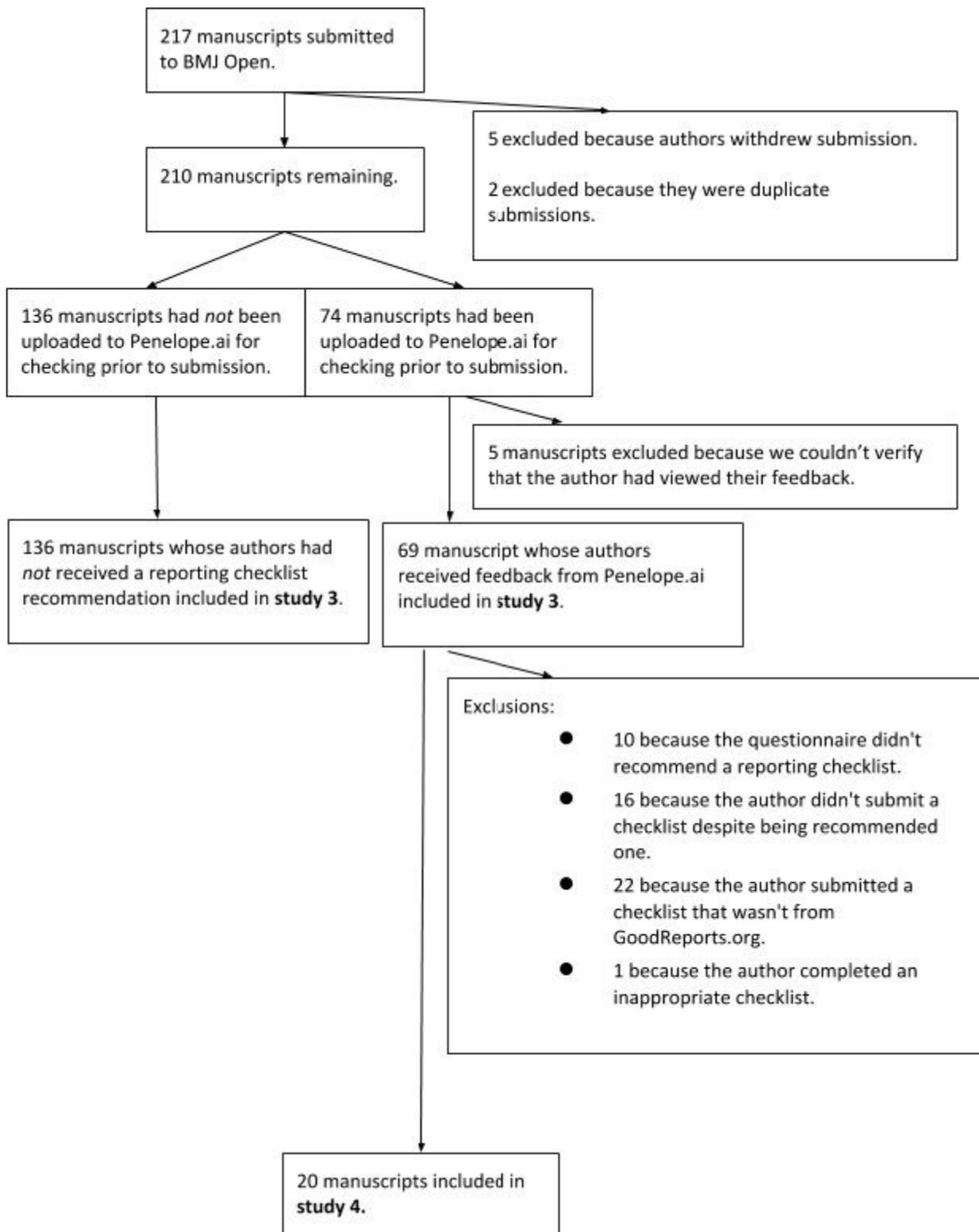
EQUATOR reporting guideline decision tree (“wizard”) developed 2016. Available from the EQUATOR website (29)





**Figure 3**

Responses to "How useful did you find the checklist" (10 = most useful)



**Figure 4**

Flow diagram showing how manuscripts were selected for rates of submission study and before-and-after study