Introduction

Table of contents

In this chapter I introduce the evidence gap I have addressed and the approach I have taken to address it. I begin by describing the prevalence and consequences of poorly reported medical research. I introduce reporting guidelines and position them as part of a complex behaviour change intervention. This intervention has had a disappointing impact, as medical research is still poorly reported. This brings me to my evidence gap: how can we get more authors to adhere to reporting guidelines? I then outline my aims, objectives, and thesis structure.

## The problem of poor reporting in health research

Facing an uncertain choice during treatment for multiple myeloma, epidemiologist Alessandro Liberati wrote “Why was I forced to make my decision knowing that information was somewhere but not available?”[1]. When I started my PhD, governments may have been asking the same question. The world was in the grip of COVID-19 and decision makers were wading through a deluge of patchy research articles missing important information [2]. In the years since, friends and family have had to make treatment decisions where the evidence is of “low certainty” because key details are missing from research articles.

A selfish silver-lining of these tumultuous years was that my family and friends finally understood the problem my thesis addresses: when medical researchers inadequately describe what they did or what they found, other people cannot understand, replicate, or use their work. Research costs huge amounts of time, money, and effort, and the written account is typically its sole legacy. When details are omitted they are lost. The remaining gaps are sources of doubt; are they accidental omissions? Oversights? Cover-ups? Whatever their source, the gaps fragment the full picture, and the potential value to patients drains away.

Early concern over reporting quality often came from frustrated reviewers unable to find the data they needed within research reports. For example, in 1963, Glick [3] found many reports of psychiatric therapy used ambiguous descriptions of treatment duration like “at least two months” or “from one to several months”. These descriptions were so vague they were “unsuitable for comparative purposes”. More recently, Dechartes found systematic reviewers could not judge the potential for bias in a third of clinical trials because of poorly described methods, thereby limiting the confidence of conclusions [4].

Reviewers are not the only people affected. When interventions are poorly described, researchers cannot appraise or repeat research. Carp [5] described how a third of 241 brain imaging studies missed information necessary to interpret and repeat them, like the number of examinations, examination duration, and the resolution of images. Doctors and service providers also need clear descriptions to replicate interventions [6]. As Feinstein noted in 1974 [7], it is difficult enough for a clinician to understand the value of unfamiliar procedure, but “it is much more difficult when he is not told what that procedure was”. For example, Davidson et al. [8] reviewed trials describing exercise interventions for chronic back pain and found authors often did not describe interventions sufficiently for other healthcare providers to copy them.

These are a mere handful of many studies documenting poor reporting in medical literature. A 2023 systematic review found 148 published between 2020-2022 alone [9]. All investigated reporting quality in different medical research disciplines, and *almost* all concluded reporting was sub-optimal. Hence, poor reporting is a long-standing problem, plagues all disciplines, devalues research, and derails the uptake of new knowledge into clinical practice.

## Reporting guidelines to the rescue?

Concern over reporting quality crescendoed through the eighties and early nineties as systematic reviews became more common. Responding to calls for “strategies”, “guides”, and “lists” to help authors prepare their manuscripts, a group of methodologists, trialists, and editors met in 1996 to create the CONsolidated Standards of Reporting Trials (CONSORT) statement [10]. CONSORT is a set of recommendations detailing what information authors should include in clinical trial reports. It comprised an article describing how it was made, a checklist, flow diagram, and (after an update in 2001) and ‘Explanation and Elaboration’ publication [11]; [12].

CONSORT proved influential, and other groups quickly developed guidelines for different research types. Reporting guidelines are like a theme and variations, where CONSORT forged a path others have followed with varying fidelity (See [Table 1](#tbl-rgs)). Most have acronym names. Most were first published as a journal article describing their development. Some, but not all, have checklists and elaboration documents. Some guideline developers publish resources as separate documents, others put them all into a single journal article. Guidelines are developed by different groups, with different composition (possibly including methodologists, editors, clinicians etc) and in different ways (e.g., some by delphi consensus). Although most follow CONSORT’s approach of presenting recommendations focussing on reporting above conduct, guidelines differ in how forceful their recommendations are and whether they also seek to influence design.

Table 1: A selection of highly cited reporting guidelines

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Guideline acronym** | **Definition** | **Applicable study type** | **Publication year** | **Development article?** | **Fillable checklist?** | **Explanatory document?** | **Other resources** | **Influences design?** |
| CONSORT | Consolidated Standards of Reporting Trials | Randomised controlled trials | 1996 #REF updated in 2001 #REF and 2010 #REF | Yes | Yes | Yes, as a separate article | Flow diagram  Website  COBWEB writing tool #REF | No |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses | Systematic Reviews and Meta-Analyses | 2009 #REF  Updated in 2021 #REF | Yes | Yes | Yes | Flow diagram  Website | No |
| ARRIVE | Animal Research: Reporting of *In Vivo* Experiments | Publications describing research involving live animals | 2010 #REF  Updated in 2020 #REF | Yes | Yes | Yes | Website  Action Plans  Compliance questionnaire | Not explicitly, but does contain design guidance |
| SRQR | Standards for Reporting Qualitative Research | Qualitative health research | 2014 [13] | Yes | No | Yes, as supplementary material that is hard to find |  |  |
| e.t.c. for all guidelines mentioned on EQUATOR’s home page  #TODO |  |  |  |  |  |  |  |  |

There are now over 500 reporting guidelines, representing the collective work of thousands of academics. The best-known guidelines are endorsed by large numbers of medical journals and the International Committee of Medical Journal Editors, and are amongst the 1% most highly cited publications indexed by Web of Science [14].

## The EQUATOR Network unites the reporting guideline movement

As reporting guidelines grew in number and the problem of poor reporting gained recognition, Doug Altman saw the need to catalogue reporting guidelines and form a community. He united academics from around the world to form The EQUATOR Network, often simply called EQUATOR, standing for Enhancing the QUAlity and Transparency Of health Research. It was the first coordinated attempt to combat poor reporting systematically and on a global scale. One of EQUATOR’s core objectives was to create a database of reporting guidelines, accessible via their website where researchers will also find training and information about developing guidelines.

## Reporting guidelines are part of a complex behaviour intervention

The publishing community began to take note. The International Committee for Medical Journal Editors encouraged journals “to ask authors to follow [reporting] guidelines” [15]. Concerned editors sought ways to adopt reporting guidelines, and more and more journals [16]; [17] have since found a range of strategies to introduce reporting guidelines into their policies, outlined in [Table 2](#tbl-journal-policies). There is variation in the degree of enforcement (from passive recommendation through to compulsory enforcement) and variation in the guidelines subject to the policy; PRISMA and CONSORT and more commonly enforced than STROBE, say, and many reporting guidelines in EQUATOR’s library are not enforced nor endorsed by any journals. Instead of listing reporting guidelines by name, many journals keep their instructions vague and merely recommend authors find an appropriate guideline on EQUATOR’s website.

Table 2: Examples of how journals have introduced reporting guidelines into their policies. Journals also differ in the reporting guidelines they enforce. For example, some journals may only have policies for randomised trials or systematic reviews, whereas other journals may enforce guidelines for other study types. To my knowledge, no journal explicitly advises *against* using a reporting guideline.

| Enforcement type | Example |
| --- | --- |
| Enforcing adherence | An editor or peer reviewer checks the article body for reporting guideline adherence and asks the author to add missing items. |
| Requesting peer reviewers use reporting guidelines | Editors ask peer reviewers to consider reporting guideline adherence as part of their review. Some editors may supply the reviewer with the relevant checklist. The reviewer can choose whether to review reporting. |
| Enforcing checklist submission | Editorial staff require authors to submit a completed reporting checklist as part of manuscript submission. Some journals may refuse to process a submission when the checklist is missing. Some journal submission systems may include fields for authors to upload their checklists, whereas other journals may expect authors to upload checklists as a supplementary file. |
| Journal endorsement | The journal’s instructions to authors recommends authors follow reporting guidelines. Guidelines may be specified, in which case journals may link to guideline specific websites, to the guideline publications, or to the EQUATOR Network website. Sometimes journals include a general statement but do not name guidelines, instead referring authors to the EQUATOR website with an instruction to follow “relevant guidance”. |
| Publisher endorsement | Sometimes reporting guideline policies are set at the level of the publisher, as is commonly done for editorial policies. Individual journals may point authors to their publisher policies. |
| No policy | Journals have no policies regarding reporting guidelines. |

Other stakeholders have begun incorporating reporting guidelines into their policies. Conferences like the Peer Review Congress ask applicants to use reporting guidelines when writing their abstracts. MedRxiv, a large preprint server, asks authors to declare they have “followed all appropriate research reporting guidelines, such as any relevant EQUATOR Network research reporting checklist(s)” [18].

EQUATOR have developed training programmes based on reporting guidelines. The training covers different ways to use reporting guidelines, including drafting manuscripts, checking manuscripts you have written, and appraising the reporting of someone else’s manuscript. Researchers have developed writing software to help authors apply reporting guidelines when drafting [19, 20], online applications to facilitate checklist and flow diagram completion [21, 22], and resources for reviewers to check compliance [23].

Hence over the years, a system has organically grown around reporting guidelines, driven by separate groups of people doing what they felt was sensible. This system includes the guidance resources themselves (the publications, checklists, flow diagrams), the websites that host those resources (guideline websites, the EQUATOR website, publisher’s websites), organisations that promote or enforse their use (EQUATOR, publishers, ICMJE), and staff at those institutions (researchers, editors, reviewers). These components all have the same aim: to influence what information researchers include in their articles.

As the Medical Research Council notes, a system with multiple components (like those listed above) is one of many hallmarks of a complex intervention: “an intervention might be considered complex because of properties of the intervention itself, such as the number of components involved; the range of behaviours targeted; expertise and skills required by those delivering and receiving the intervention; the number of groups, settings, or levels targeted; or the permitted level of flexibility of the intervention or its components.”. The reporting guideline system exhibits these sources of complexity, as described in [Table 3](#tbl-complexity).

Table 3: Sources of complexity within reporting guidelines and the system drives their use.

| Source of complexity | Example |
| --- | --- |
| Number of components involved | Reporting guidelines often consist of guidance documents, checklists, and flow diagrams, and other tools. These are disseminated through websites, publishing platforms, submission systems, and they are endorsed and enforced by staff at stakeholders including publishers, the EQUATOR Network, conference organisers, and pre-print platforms. |
| Range of behaviours targeted | Guidelines comprise “reporting items”. Some items are relatively simple, like asking the author to specify their study design in the title. Others are harder, perhaps because they require time, expertise, or prerequisite tasks. For instance, some items may require authors to have conducted their study or analysis in a certain way, or to have collected particular information. |
| Expertise and skills required by those delivering and receiving the intervention | Academics from a particular field write reporting guidelines for their peers (as opposed to a lay audience), and so authors, editors, and reviewers must have sufficient expertise to use them. |
| The number of groups, settings, or levels targeted | Groups: Users of reporting guidelines differ in their field of expertise, their experience, place of work.  Settings: Although mostly written with authoring in mind, most guideline developers may also hope their resources are used by editors or peer reviewers for checking or appraising research articles.  Guidelines are written with individuals in mind, but their efficacy is generally measured at group level (e.g. articles from a particular field published in a period time). |
| Flexibility of the intervention or its components | There is variation between guideline content, resources, and the implementation strategies that development groups, publishers, and other stakeholders employ. |

Viewing reporting guidelines as part of a complex behaviour change intervention may seem novel. Researchers often call reporting guidelines “tools” or “strategies” (#REF). In “A history of the EQUATOR Network”, Doug Altman refers to reporting guidelines as “resources” that “influence” reporting, but he does not call them interventions. However, I believe my perspective is not radical. I will now outline studies exploring the efficacy of reporting guidelines, and argue that these studies take a systems perspective too although they seldom acknowledge it explicitly.

## Reporting quality has improved over time but remains sub optimal

### Evidence from observational studies

Many studies have compared quality of reporting before and after reporting guidelines were published and/or journals began asking authors to use them. For example, in 2023 Kilicoglu et al. [24] analysed 176,620 randomized trial reports published between 1966 and 2018. They used software to identify sentences pertaining to CONSORT methodology items. They found reporting quality in clinical trials has improved over time but remains sub optimal: articles published before 1990 reported 24% of CONSORT items, whereas articles published between 2010-2018 reported 48%. The fastest improvement occurred in the years immediately after CONSORT was published.

This result mirrors similar, manual, efforts to monitor reporting quality over time. For example, Dechartres et al. [4] looked at CONSORT items related to bias in 20,920 trial reports. They found information on sequence generation was missing from 69% of articles published before CONSORT (1986-1990), but only 31% of articles published after CONSORT (2011-2014). The proportion of articles missing information on allocation concealment fell from 70% to 45% over the same periods.

These two examples focus on clinical trials, but other research types have seen modest improvements too. De Jong et al. [25] reviewed 284 qualitative evidence syntheses. These syntheses appraised whether their primary studies had adhered to COREQ (a reporting guidelines for qualitative research). Studies published before COREQ reported 16 of COREQ’s 32 items, whereas studies published after reported 18. Similarly, when comparing the reporting quality of genetic association studies, Nedovic et al. [26] found reporting quality was better in generic association articles published after STREGA but only in journals that endorsed STREGA (63% vs 49% showing full adherence).

Not all studies have found relationships like these. For example, Howell et al. found no improvement in the reporting of quality improvement studies after the SQUIRE guidelines were published [27]. Similarly, Pouwels et al. found no improvement in observational epidemiology following STROBE’s publication [28] and another study [29] found that reporting quality improved *before* STROBE was published, but not *afterwards*.

Other studies have explored the effect of different journal policies. For example, Hopewell et al. [30] assessed the reporting of clinical trial abstracts before and after the publication of CONSORT for Abstracts and compared journals that a) had no reporting guideline policy b) endorsed the guideline and c) actively enforced guideline adherence. They only found an effect in the active enforcement group, where 5.41 items (out of 9) were reported, which was 50% higher than expected. No improvement was seen in journals that endorsed the guideline without enforcing it.

Some observational studies have focussed on a single time period. In 2018, before my PhD, I collaborated with EQUATOR and BMJ Open to compare manuscript versions before and after authors completed a reporting checklist [22]. Authors made few edits to their work. Of 20 included authors, three added information for a single reporting item, one added information for two items, and one added information for six items. The remaining 15 authors made no changes. On average, manuscripts described 57% of necessary reporting items before the author completed the checklist and 60% afterwards.

In 2018 the senior managing editor of the Journal of the National Cancer Institute asked 2000 submitting authors whether they had used a reporting guideline [31]. She then asked peer reviewers to rate manuscripts for their clarity and adherence to reporting guidelines. Declared guideline use was associated with better adherence to guidelines, but not associated with improved clarity nor acceptance rates.

### Evidence from intervention studies

Some experimental studies have tried to isolate the effect of journal policies by randomising authors. In 2015 PLOS One randomly allocated 1689 incoming submissions reporting *in vivo* animal research manuscripts to either a) request completion the ARRIVE reporting checklist or b) usual practice [32]. No article achieved full compliance with ARRIVE, and only one sub item (details of animal husbandry) showed improvement between groups.

In another study [33], 197 authors submitting to 46 participating journals were randomly allocated to receive either a) access to an online tool (WebCONSORT) to generate customised reporting checklists and flow diagrams based on CONSORT and its extensions or b) a standard CONSORT flow diagram generator without a reporting checklist. There was no difference in reporting quality between groups: authors only reported half of required items.

In two earlier studies, Cobo et al. explored the roll of peer review. Simply providing peer reviewers with reporting guidelines had no effect [34]. Adding a reviewer whose task was to check for guideline adherence did lead to improved reporting quality [35], but “the observed effect was smaller than hypothesised and not definitively demonstrated”.

It would be unsurprising to find that the stricter the enforcement, the better the adherence. Pandis et al. [36] describe an enforcement strategy used in a small Dentistry journal where the associate editor checked manuscripts for adherence to CONSORT. The associate editor would complete a CONSORT checklist for each manuscript, making note of unclear or unreported items and suggesting ways to improve the manuscript. This would be sent back to the author. Resubmitted manusctripts were subject to the same process, and the manuscript would only be sent out for peer review once the reporting was deemed satisfactory. Over two years, 23 manuscripts were handled in this way. The policy was effective. All studies reported at least 33 of 37 CONSORT items (compared to 15 items before the policy was introduced). However, even with this heavy handed approach, “four items were still unreported in all trials: changes to methods (3b), changes to outcomes (6b), interim analysis (7b), and trial stopping (14b).”.

To summarise, reporting standards may have improved over the last two decades. There is some evidence that reporting guidelines may have contributed to this change, but if they have, their effect has only been modest and the bottom line is that most research still does not include the details these guidelines call for. There is huge room for improvement. In a systematic review of 124 studies assessing adherence to one of eight reporting guidelines, Jin et al [37] found 88% of studies reported suboptimal adherence to reporting guidelines. Similarly, Dal Santo et al [38] reviewed 148 studies of reporting quality from the previous few years and almost all described reporting quality as suboptimal.

### Studies exploring the efficacy of reporting guidelines are actually exploring the efficacy of the reporting guideline *system*

Because reporting guidelines do not exist in a vacuum, it is difficult to separate the guidelines themselves from the policies, people, websites, and tools involved in their implementation. For example, many before-and-after studies use the publication of the guideline as their defining time point. However, it is impossible for these studies to disentangle the effect of guidelines coming into existence with the effect of subsequent journal policies and editorial practices. Similarly, in experimental studies comparing the effect of asking authors to complete a checklist or use a resource, the intervention groups included changes to editorial workflows. These changes were external to the resource being tested, but could be equally important to its success: in the WebCONSORT study, editors’ inability to identify randomised trial reports was an important source of failure external to the tool being tested. Hence, in describing reporting guidelines as being part of a complex behaviour change intervention, I believe I am explicitly articulating a systems perspective that previous studies have hinted at, and I am exploring that system’s scope in more granularity.

## Evidence gap: What more could be done to improve guideline adherence?

Some of the articles I have cited end with rallying cries like “major improvements need active enforcement” [39]. It would be tempting to look at Pandis’ results as support for heavy editorial enforcement being the best option, but this approach may not generalise to other journals and other guidelines. The dentistry journal in this study was small. Only 23 manuscripts underwent this treatment over 2 years, and despite giving “30 to 60 minutes” of editorial attention to each manuscript, not all completely adhered to CONSORT. The study authors admit the benefits should be “considered in the light of the additional time requirement and need for greater editorial input during the peer review process”.

Other articles have called for lighter forms of enforcement; “We need to promote more active implementation, such as submission of the checklist with the manuscript” wrote Dechartres [4], and “it is not sufficient for journals to simply recommend the use of STREGA to authors in the authors’ instructions; instead, journals should require submission of the STREGA checklist together with the manuscript” wrote Nedovic [26]. But the PLOS One [32] and BMJ Open studies [22] found little effect of checklist completion on reporting quality. Additionally, the PLOS One study found that enforcing checklists, although less burdensome than the editorial enforcement described by Pandis, still came with costs to both editors and authors and significantly prolonged publication times. Peer reviews focussing on reporting might help [35] but is relatively unexplored as an option.

These studies all focussed on different methods of enforcement. Some include incidental findings hinting at areas-for-improvement unfixable by enforcement alone. For example, because reviewers assessing adherence in the PLOS One study did not always agree or fully understand the guidance, the study authors suggested refining the guideline’s “content” and “perceived clarity”. In the WebCONSORT study, [33] Hopewell et al. made some guesses for why their intervention failed. They had to exclude 39% of manuscripts because editors had incorrectly identified them as randomised trials, and a quarter of authors selected inappropriate extensions. As a solution, they suggest a tool to help authors and editors identify study types. The study authors also raised other hypotheses to explain why their intervention failed: perhaps the custom combined checklists were too long, unclear, or perhaps giving feedback during manuscript revision was too late.

Both of these studies may have benefited from a qualitative component to understand why the interventions were not working. In our BMJ Open study we surveyed authors after they completed checklists. Many reported finding the checklist too long, confusing, or irrelevant. However, because we used a multiple choice question with a (small) box for a free text answer, and because we did not survey authors if they *did not* complete a checklist, authors may have faced other barriers too.

These studies suggest authors may face barriers when trying to use reporting guidelines that require solutions beyond enforcement. Noting the outcome assessors’ confusion in the PLOS One study, ARRIVE’s developers took steps to refine its clarity when they revised it (#REF). They also decided to prioritise items to make the guidance quicker to apply. Hopewell et al. created WebCONSORT because they worried combining CONSORT with its extensions may be “cumbersome and difficult” without providing evidence for this claim. In 2018 I worked with EQUATOR as a freelance developer to create GoodReports.org, a website where authors could find and fill out checklists on-line. The website addressed two barriers. Firstly a questionnaire helped authors find the right guideline. Secondly, authors could complete checklists easily (previously, some reporting guidelines came with uneditable PDF checklists).

These innovation efforts shared limitations. None took steps to identify barriers thoroughly. By focussing on a few barriers they may have neglected others or introduced new ones. For example, in trying to make combining checklists easier, WebCONSORT may inadvertently made checklists longer, and increased the risk of authors selecting inappropriate guidance. Secondly, these studies did not systematically consider options to solve those barriers. For example, ARRIVE’s development team decided to prioritize items as a way to make the guidance quicker to apply, but this is not the only solution. They could also have considered reducing the number of items, making guidance more concise, providing suggested wording or creating tools to speed up writing. Thirdly, although the studies describe their innovations, they do not always describe changes beyond the tool in question. For example, implementing WebCONSORT and PLOS One’s checklist policy also involved changes in editorial practice. Nor do these studies always clearly explain how changes are expected to alter behaviour.

## Summary

In summary, I’ve explained that a lot of medical research is poorly reported, and that this makes it difficult for other researchers to understand, appraise, synthesize, or replicate studies. This, in turn, makes research less useful to patients.

I’ve introduced reporting guidelines, created by the research community with the aim to improve reporting quality. I’ve described the system of tools, websites, people, and policy that has organically grown around reporting guidelines, and I have argued that this system forms a complex behaviour change intervention with the goal of altering what authors write.

I’ve discussed how this system has had only a modest effect on reporting quality, at best. I’ve described how studies exploring modifications to this system are limited because they did not explore barriers thoroughly, and lacked a systematic method to identify options to address those barriers.

## Aims and Objectives

My aim was to identify and address barriers preventing authors from adhering to reporting guidelines. I wanted to explore the entire reporting guideline system, and I wanted to be thorough: I wanted to identify as many barriers as possible, and as many solutions as possible, before deciding which to implement.

My objectives were:

* To identify factors that may limit reporting guideline impact by synthesising existing research and by evaluating the EQUATOR Network website (addressed in chapters 3)
* To work with key stakeholders to identify intervention changes to address these limiting factors (addressed in chapters 6)
* To implement these changes (described in 10)
* To refine the new intervention in response to feedback from authors (addressed in chapter 11)

My thesis bares many of the hallmarks of pragmatism. I used both qualitative and quantitative methods. Constraints (like time and access to participants) influenced my decisions. I balanced participants’ views with my own; I sought to remove my views as much as possible in all chapters except for the workshops I conducted with EQUATOR (chapter 7) and when designing the intervention (chapter 10). I balanced inductive and deductive reasoning; my early chapters were exploratory and inductive, and my later chapters became increasingly deductive as my focus narrowed and I relied more heavily on a framework.

## Thesis structure

**Chapter 2 - Reflections on starting my DPhil**

I reflect on my background and my prior held opinions, and those of my supervision team, and how these may have influenced the direction of this thesis.

**Chapter 3 - What facilitators and barriers might researchers encounter when using reporting guidelines? Part 1: A thematic synthesis**

The next three chapters pertain to my first objective - to identify possible reasons as to why reporting guidelines have had only a limited impact on reporting quality. This chapter describes a thematic synthesis of studies that qualitatively explored authors’ experiences of using reporting guidelines, where I sought to identify what may influence whether an author successfully adheres to reporting guidance.

**Chapter 4 - What facilitators and barriers might researchers encounter when using reporting guidelines? Part 2: Describing the questions asked in quantitative surveys.**

This chapter builds on the previous one by identifying additional possible influences from the content of quantitative survey questions.

**Chapter 5 - A service evaluation of equator-network.org**

This chapter describes a service evaluation of the EQUATOR Network website which, although an important piece of the reporting guideline infrastructure, was rarely explored by studies reviewed in the previous two chapters. From this evaluation, I then infer possible barriers that authors may encounter when trying to find and access reporting guidelines from EQUATOR’s website.

**Chapter 6 - Selecting the Behaviour Change Wheel framework**

The next 4 chapters pertain to my second objective - identifying intervention changes. This chapter introduces the Behaviour Change Wheel, which is a framework for designing and defining behaviour change interventions. I explain how my thesis gained form at this point in time; my view of reporting guidelines as a system crystallised, and Charlotte Albury joined my supervision team as my plans took an unexpected qualitative turn. I wanted my thesis to accurately reflect the twists and turns of my DPhil, and so I introduce my chosen framework in this middle chapter instead of the introduction which may be more customary.

**Chapter 7 - Following the BCW Guide: Workshops with EQUATOR**

This chapter describes how I lead workshops with UK EQUATOR center staff to identify intervention options using Behaviour Change Wheel framework.

**Chapter 8 - Generating ideas to address factors limiting reporting guideline impact: workshops with EQUATOR and focus groups with developers, publishers, and experts**

This chapter reports focus groups where I collected ideas on how intervention options could be realised.

**Chapter 9 - Defining Intervention Content**

In this chapter, I bring together the outputs of the previous two chapters to create a table of intervention components.

**Chapter 10 - Developing intervention components into a prototype**

This chapter concerns my third objective; implementing the intervention changes by redesigning a reporting guideline (SRQR) and the EQUATOR Network website’s home page.

**Chapter 11 - Refining the intervention: qualitative study with authors**

In this chapter I address my final objective by refining the intervention in response to feedback from authors. I describe a qualitative study where I used observation, think aloud, structured interviews, and a writing evaluation, to gather feedback from an international sample of authors.

1. Liberati A (2004) [An unfinished trip through uncertainties](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC351869). BMJ : British Medical Journal 328:531

2. Ziemann S, Paetzolt I, Grüßer L, Coburn M, Rossaint R, Kowark A (2022) [Poor reporting quality of observational clinical studies comparing treatments of COVID-19 a retrospective cross-sectional study](https://doi.org/10.1186/s12874-021-01501-9). BMC Medical Research Methodology 22:23

3. Glick BS (1963) [Inadequacies in the reporting of clinical drug research](https://doi.org/10.1007/BF01562195). The Psychiatric Quarterly 37:234–244

4. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, Perrodeau E, Altman DG, Ravaud P (2017) [Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: Research on research study](https://doi.org/10.1136/bmj.j2490). BMJ 357:j2490

5. Carp J (2012) [The secret lives of experiments: Methods reporting in the fMRI literature](https://doi.org/10.1016/j.neuroimage.2012.07.004). NeuroImage 63:289–300

6. Glasziou P, Meats E, Heneghan C, Shepperd S (2008) [What is missing from descriptions of treatment in trials and reviews?](https://doi.org/10.1136/bmj.39590.732037.47) BMJ 336:1472–1474

7. Feinstein AR (1974) [Clinical biostatistics. XXV. A survey of the statistical procedures in general medical journals](https://doi.org/10.1002/cpt197415197). Clinical Pharmacology and Therapeutics 15:97–107

8. Davidson SRE, Kamper SJ, Haskins R, Robson E, Gleadhill C, da Silva PV, Williams A, Yu Z, Williams CM (2021) [Exercise interventions for low back pain are poorly reported: A systematic review](https://doi.org/10.1016/j.jclinepi.2021.05.020). Journal of Clinical Epidemiology 139:279–286

9. Santo TD, Rice DB, Amiri LSN, Tasleem A, Li K, Boruff JT, Geoffroy M-C, Benedetti A, Thombs BD (2023) [Methods and results of studies on reporting guideline adherence are poorly reported: A meta-research study](https://doi.org/10.1016/j.jclinepi.2023.05.017). Journal of Clinical Epidemiology 159:225–234

10. Begg C, Cho M, Eastwood S, et al (1996) [Improving the Quality of Reporting of Randomized Controlled Trials: The CONSORT Statement](https://doi.org/10.1001/jama.1996.03540080059030). JAMA 276:637–639

11. Altman DG (1996) [Better reporting of randomised controlled trials: The CONSORT statement](https://doi.org/10.1136/bmj.313.7057.570). BMJ (Clinical research ed) 313:570–571

12. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T (2001) [The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration](https://doi.org/10.7326/0003-4819-134-8-200104170-00012). Annals of Internal Medicine 134:663–694

13. O’Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA (2014) [Standards for Reporting Qualitative Research: A Synthesis of Recommendations](https://doi.org/10.1097/ACM.0000000000000388). Academic Medicine 89:1245–1251

14. Caulley L, Cheng W, Catalá-López F, Whelan J, Khoury M, Ferraro J, Husereau D, Altman DG, Moher D (2020) [Citation impact was highly variable for reporting guidelines of health research: A citation analysis](https://doi.org/10.1016/j.jclinepi.2020.07.013). Journal of Clinical Epidemiology 127:96–104

15. ICMJE | Recommendations | Preparing a Manuscript for Submission to a Medical Journal.

16. Koch M, Riss P, Umek W, Hanzal E (2016) [The explicit mentioning of reporting guidelines in urogynecology journals in 2013: A bibliometric study](https://doi.org/10.1002/nau.22726). Neurourology and Urodynamics 35:412–416

17. Sharp MK, Tokalić R, Gómez G, Wager E, Altman DG, Hren D (2019) A cross-sectional bibliometric study showed suboptimal journal endorsement rates of STROBE and its extensions. Journal of clinical epidemiology 107:42–50

18. medRxiv.org - the preprint server for Health Sciences.

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

20. Hawwash D, Sharp MK, Argaw A, Kolsteren P, Lachat C (2019) Usefulness of applying research reporting guidelines as Writing Aid software: A crossover randomised controlled trial. BMJ Open. <https://doi.org/10.1136/bmjopen-2019-030943>

21. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA (2022) [PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis](https://doi.org/10.1002/cl2.1230). Campbell Systematic Reviews 18:e1230

22. Struthers C, Harwood J, de Beyer JA, Dhiman P, Logullo P, Schlüssel M (2021) [GoodReports: Developing a website to help health researchers find and use reporting guidelines](https://doi.org/10.1186/s12874-021-01402-x). BMC medical research methodology 21:217

23. Compliance Questionnaire | ARRIVE Guidelines.

24. Kilicoglu H, Jiang L, Hoang L, Mayo-Wilson E, Vinkers CH, Otte WM (2023) Methodology reporting improved over time in 176,469 randomized controlled trials. Journal of Clinical Epidemiology. <https://doi.org/10.1016/j.jclinepi.2023.08.004>

25. de Jong Y, van der Willik EM, Milders J, Voorend CGN, Morton RL, Dekker FW, Meuleman Y, van Diepen M (2021) [A meta-review demonstrates improved reporting quality of qualitative reviews following the publication of COREQ- and ENTREQ-checklists, regardless of modest uptake](https://doi.org/10.1186/s12874-021-01363-1). BMC Medical Research Methodology 21:184

26. Nedovic D, Panic N, Pastorino R, Ricciardi W, Boccia S (2016) [Evaluation of the Endorsement of the STrengthening the REporting of Genetic Association Studies (STREGA) Statement on the Reporting Quality of Published Genetic Association Studies](https://doi.org/10.2188/jea.JE20150173). Journal of Epidemiology 26:399–404

27. Howell V, Schwartz AE, O’Leary JD, Donnell CM (2015) [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](https://doi.org/10.1136/bmjqs-2014-003737). BMJ Quality & Safety 24:400–406

28. Pouwels KB, Widyakusuma NN, Groenwold RHH, Hak E (2016) [Quality of reporting of confounding remained suboptimal after the STROBE guideline](https://doi.org/10.1016/j.jclinepi.2015.08.009). Journal of Clinical Epidemiology 69:217–224

29. Bastuji-Garin S, Sbidian E, Gaudy-Marqueste C, Ferrat E, Roujeau J-C, Richard M-A, Canoui-Poitrine F, European Dermatology Network (EDEN) (2013) [Impact of STROBE statement publication on quality of observational study reporting: Interrupted time series versus before-after analysis](https://doi.org/10.1371/journal.pone.0064733). PloS One 8:e64733

30. Hopewell S, Ravaud P, Baron G, Boutron I (2012) [Effect of editors’ implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: Interrupted time series analysis](https://doi.org/10.1136/bmj.e4178). BMJ 344:e4178

31. Botos J (2018) [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](https://doi.org/10.1186/s41073-018-0052-4). Research Integrity and Peer Review 3:7

32. Hair K, Macleod MR, Sena ES, et al (2019) [A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines (IICARus)](https://doi.org/10.1186/s41073-019-0069-3). Research Integrity and Peer Review 4:12

33. Hopewell S, Boutron I, Altman DG, et al (2016) [Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: Results of a randomised controlled trial](https://doi.org/10.1186/s12916-016-0736-x). BMC Medicine 14:199

34. Cobo E, Selva-O’Callagham A, Ribera J-M, Cardellach F, Dominguez R, Vilardell M (2007) [Statistical Reviewers Improve Reporting in Biomedical Articles: A Randomized Trial](https://doi.org/10.1371/journal.pone.0000332). PLoS ONE 2:e332

35. Cobo E, Cortés J, Ribera JM, et al (2011) [Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: Masked randomised trial](https://doi.org/10.1136/bmj.d6783). BMJ 343:d6783

36. Pandis N, Shamseer L, Kokich VG, Fleming PS, Moher D (2014) [Active implementation strategy of CONSORT adherence by a dental specialty journal improved randomized clinical trial reporting](https://doi.org/10.1016/j.jclinepi.2014.04.001). Journal of Clinical Epidemiology 67:1044–1048

37. Jin Y, Sanger N, Shams I, et al (2018) [Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? A systematic review of reviews: An update](https://doi.org/10.2147/JMDH.S155103). Journal of Multidisciplinary Healthcare 11:495–510

38. Santo TD, Rice DB, Amiri LSN, Tasleem A, Li K, Boruff JT, Geoffroy M-C, Benedetti A, Thombs BD (2023) [Methods and results of studies on reporting guideline adherence are poorly reported: A meta-research study](https://doi.org/10.1016/j.jclinepi.2023.05.017). Journal of Clinical Epidemiology 159:225–234

39. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Moher D, Wager E (2014) [Reducing waste from incomplete or unusable reports of biomedical research](https://doi.org/10.1016/S0140-6736(13)62228-X). Lancet (London, England) 383:267–276