# Addressing harms of screening – A review of outcomes in Cochrane reviews and suggestions for next steps

Johansson, Minna; Borys, Franciszek; Peterson, Hanna; Bilamour, Giulia; Bruschettini, Matteo; Jørgensen, Karsten Juhl

*Published in:*
Journal of Clinical Epidemiology

*DOI:*
10.1016/j.jclinepi.2020.09.030

*Publication date:*
2021

*Document version:*
Final published version

Journal of Clinical Epidemiology 129 (2021) 68−73

Journal of Clinical Epidemiology

REVIEW

# Addressing harms of screening − A review of outcomes in Cochrane reviews and suggestions for next steps

Minna Johansson[a,b,*], Franciszek Borys[b], Hanna Peterson[b], Giulia Bilamour[b], Matteo Bruschettini[b], Karsten Juhl Jørgensen[c]

[a]*Cochrane Sustainable Healthcare, Hedegatan 38, Uddevalla 45152, Sweden*
[b]*Cochrane Sweden, Skåne University Hospital, Wigerthuset, Remissgatan 4, Lund 222 42, Sweden*
[c]*The Nordic Cochrane Centre, Rigshospitalet Dept. 7811, Blegdamsvej 9, Copenhagen DK-2100, Denmark*

Accepted 21 September 2020; Published online 1 October 2020

## Abstract

**Objective:** To investigate if Cochrane reviews that assess screening interventions address their major harms.

**Study design and setting:** A systematic search for Cochrane reviews that assess screening interventions was performed. Two authors independently screened abstracts, assessed full-texts, and extracted data from included reviews. For each review, two authors judged whether each predefined harm was relevant. When the harm was judged as of questionable relevance, the review was excluded from the denominator in our calculations.

**Results:** Forty-seven reviews were included. Overdiagnosis was addressed in 6 of 39 (15%), overtreatment in 7 of 43 (16%), and psychosocial consequences in 30 of 47 (64%) of reviews where this was judged relevant. When data on harms were included, they were generally not treated with the same methodological rigor as the benefits, with no assessment of the risk of bias or certainty of the evidence. About half of the Abstracts, Plain Language Summaries, and Summary of Findings tables did not include any harms.

**Conclusion:** The underreporting of harms of screening in Cochrane reviews likely reflects primary research and is problematic. We call for broad collaboration to develop reporting guidelines and core outcome sets for studies of screening interventions. 

*Keywords:* Screening; Research reporting; Harm; Overdiagnosis; Overtreatment

## 1. Background

The main objective of screening programs is to improve prognosis by earlier detection of disease, precursors of disease, or risk factors for disease. In general, the most important beneficial effect of screening is a decrease in mortality and/or morbidity from the condition screened for [1−3]. However, there are also harms associated with screening people who have no symptoms of the disease. These include false positive and false negative findings, overdiagnosis and overtreatment of harmless conditions, and their physical and psychosocial consequences [1−3]. As for all other types of medical interventions, patients, clinicians, and policymakers must be able to balance benefits against harms to judge whether a screening intervention is worthwhile or not [1−3]. Therefore, solid evidence of the magnitude of both benefits and harms is needed, as well as an assessment of the certainty of that evidence. However, the harms of screening may not be adequately investigated in primary research; a systematic review found that overdiagnosis was

Prospero protocol registration number: PROSPERO 2019 CRD42019138110.

Role of funding source: This study was funded by Cochrane Sweden and the Nordic Cochrane Center. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; the decision to submit the manuscript for publication.

Declaration of interest: MJ, MB, and KJJ all hold positions within Cochrane. We declare no other conflicts of interest.

Transparency statement: MJ affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; that discrepancies, if any, from the study as originally planned have been explained.

Ethics approval: Not applicable.

Data sharing: All extracted data is available in the Appendix.

This research or the matters expressed in this paper do not represent the views of Cochrane.

* Corresponding author. Tel.: 0046-733106170.
*E-mail address:* minna.johansson@vgregion.se (M. Johansson).

https://doi.org/10.1016/j.jclinepi.2020.09.030

**What is new?**

- Harms of screening are often not included in Cochrane reviews and protocols of screening interventions. When data on harms are reported in reviews, they are generally not treated with the same methodological rigor as the benefits.
- The underreporting of harms of screening in Cochrane reviews likely reflects primary research and is problematic since patients, clinicians, and policymakers must be able to balance benefits against harms to judge whether a screening intervention is worthwhile or not.
- A broad collaboration to develop reporting guidelines and core outcome sets that take harms and benefits of screening interventions into consideration is needed.

quantified in only 7% of randomized trials of cancer screening [4], and another systematic review found that psychosocial harms were not adequately evaluated for any of the five screening programs investigated [5]. Underreporting of harms in primary studies is not specific to screening but well documented for many medical interventions [6–8]. There is an extensive literature on the specific methodological challenges posed by trying to address the harms of interventions in evidence synthesis [6–8].

Even if the primary studies do not adequately include or quantify the harms of screening, it is still important to assess the evidence base. First, this helps to clarify gaps in our knowledge, which is important information for patients, clinicians, and policymakers to consider when making a decision. Second, clarification of such gaps in knowledge can help outline a direction for future research, including outcomes that are important to report in future primary studies. The aim of this review was to investigate if Cochrane reviews and protocols that assess screening interventions address the major harms of screening.

## 2. Methods

We performed a systematic search for Cochrane reviews and protocols that assess screening interventions and investigated whether these addressed major harms of screening or not.

A protocol for this review is published on PROSPERO and includes a set of predefined harms outcomes [9]. This protocol was submitted before our search and data extraction begun, but due to delays in the editorial process, it was not published until this was completed. No changes were made to the protocol between submission and publication. Mail correspondence with PROSPERO is available on request.

We searched the Cochrane Library on June 5, 2019, and updated the search on July 26, 2020, for all Cochrane reviews and protocols that had the word ‘‘screening’’ in the title or as a keyword. We included all Cochrane reviews and protocols investigating the effects of any screening intervention compared to no screening or compared to another screening intervention, which aimed to improve mortality and/or morbidity from any single disease, precursor, or risk factor. We included protocols when the corresponding Cochrane review was not yet published. Reviews on antenatal genetic screening were not included, since benefits and harms, as well as ethical considerations of such screening interventions, differ importantly from other types of screening. Reviews on other types of antenatal screening (for example, screening for gestational diabetes) were included. Reviews of interventions to increase screening uptake or informed choice in screening, as well as diagnostic test accuracy reviews, were not included.

Two authors independently screened titles and abstracts and assessed full-text copies published in the Cochrane Library of potentially relevant reviews and protocols for inclusion (FB, MJ). For each review, two authors independently extracted data from included reviews and protocols into a data extraction form (FB, GB, HP).

We assessed whether the Cochrane reviews and protocols reported, or planned to report, the outcomes listed below (primary and secondary refers to our predefined priority of outcomes, not that reported in the review). The outcomes were considered ‘‘reported’’ if they were included as outcomes in the Methods section or if they were reported in the Result section (i.e., even if they were not prespecified as an outcome in the Methods section). We noted only whether the authors looked for these outcomes or not, i.e., we did not take into account whether they actually found any data for the outcomes or not.

Primary:

- overdiagnosis
- overtreatment
- psychosocial consequences (of for example labeling, the screening test itself, follow-up procedures or treatment due to screening)

Secondary:

- false-positive findings
- false-negative findings
- incidental findings
- somatic complications (due to, for example, the screening test itself, follow-up procedures or treatment due to screening)
- additional invasive diagnostic procedures due to screening
- all-cause mortality (this outcome was not defined as a benefit or harm as this depends on the combined direction of effect of both benefit and harm)

**Table 1.** Reporting of harms of screening in Cochrane reviews and protocols of screening interventions

| Outcome | No of reviews/protocols where the outcome was judged to be relevant | No of reviews/protocols that reported on the outcome | No of reviews/protocols that did not report on the outcome | Percentage of reviews/protocols that reported on the outcome of those where it was judged to be relevant |
|---|---|---|---|---|
| Overdiagnosis | 39 | 6 | 33 | 15% |
| Overtreatment | 43 | 7 | 36 | 16% |
| Psychosocial consequences | 47 | 30 | 17 | 64% |
| False positives | 45 | 17 | 27 | 38% |
| False negatives | 45 | 14 | 30 | 31% |
| Incidental findings | 27 | 2 | 25 | 7% |
| Somatic complications | 43 | 19 | 24 | 44% |
| Additional invasive diagnostic procedures | 27 | 9 | 18 | 33% |
| All-cause mortality | 32 | 29 | 3 | 91% |

These outcomes were chosen based on a framework for harms of screening proposed by Harris and colleagues [1]. For each type of screening, two authors (MJ, KJJ) judged whether each harm outcome could be relevant to report or not (judgments available in Appendix). Those reviews, where harm outcomes were judged to be of questionable relevance, were excluded from the denominator in our calculations.

We also planned to investigate whether treatment burden (for the participant), opportunity costs in terms of human resources (health care personnel), and opportunity costs in terms of financial resources (for the health care system), were considered in the Discussion or Conclusion sections of the reviews. However, we were not able to perform this analysis as planned (see Departures from protocol in Discussion).

## 3. Results

Our search strategy identified 94 reviews and protocols that we screened for eligibility based on the title and abstract. Forty-one reviews and protocols were excluded, and 53 reviews and protocols were assessed for full-text eligibility. Of these, 6 reviews and protocols were excluded (one review was on overdiagnosis only, one review was on a test instigated due to symptoms (i.e., not screening), two protocols had been converted to new review formats, and two protocols had later been published as included full reviews). Thus, 47 reviews and protocols were included (43 reviews and 4 protocols).

The harms of screening were often not reported in Cochrane reviews and protocols of screening interventions (Table 1). Overdiagnosis was reported in 6 of 39 (15%) of the reviews and protocols where it was judged to be relevant, and overtreatment was reported in 7 of 43 (67%) of the reviews and protocols. Psychosocial consequences were reported more often; in 30 of 47 (64%) of the reviews and protocols where it was judged to be relevant. However, data on psychosocial consequences were often presented as reported in the original study, without attempts to perform meta-analyses, nor were adequate assessment of the risk of bias or the certainty of the evidence presented for this outcome. Also, when other harm outcomes were assessed, they were generally not treated with the same methodological rigor as the benefits. For our secondary harm outcomes, reporting ranged from 7 to 44% (Table 1). Eleven of 47 reviews and protocols (23%) included harm as a primary outcome, while the remaining 77% did not.

In the Summary of Findings tables, an average of 3.0 benefits and 0.7 harms were reported, and 12 of 18 (67%) of all Summary of Findings tables did not include any harms (Table 2). In the Abstracts, an average of 2.8 benefits and 1.2 harms were mentioned, and 18 of 43 (42%) did not mention any harm (Table 2). In the Plain Language Summaries, an average of 2.5 benefits and 0.8 harms were mentioned, and 25 of 43 (58%) did not mention any harm (Table 2).

We had planned to investigate whether treatment burden and opportunity costs in terms of financial and human

**Table 2.** Reporting of benefits and harms of screening in Abstracts, Plain Language Summaries and Summary of Findings tables of Cochrane reviews and protocols of screening interventions

| Review section | No. of benefits reported (mean) | No. of harms reported (mean) | Percentage where no benefit was reported | Percentage where no harm was reported |
|---|---|---|---|---|
| Abstract | 2.8 | 1.2 | 0% | 42% |
| Plain language summary | 2.5 | 0.8 | 2% | 58% |
| Summary of findings table | 3.0 | 0.7 | 0% | 67% |

resources were considered in the Discussion or Conclusion section of the reviews. We found that treatment burden and opportunity costs in terms of human resources within the healthcare system were rarely considered, while financial costs were more often considered in the discussion, as well as included as an outcome. However, the reporting was too inconsistent and unclear for us to arrive at any meaningful numbers with regard to these outcomes (see Discussion).

## 4. Discussion

In this systematic review, we found that major harms of screening were often not reported in Cochrane reviews and protocols of screening interventions, even when specifically assessed as relevant for that type of screening. Overdiagnosis, often considered the most important harm of screening [1,10], was addressed in 15% of the reviews and protocols, and overtreatment in 16%. The psychosocial consequences of screening were considered more often; in 64% of the reviews and protocols. However, the data on the psychosocial consequences were most often simply presented as reported in the studies, without summarizing the results or adequate assessment of the associated risk of bias or the certainty of the evidence. This may be especially problematic for these outcomes since the methodological quality of studies of the psychosocial consequences of screening is often poor [5]. For other harm outcomes, when they were included, these were generally also not treated with the same methodological rigor as beneficial outcomes, i.e., meta-analyses were not performed, and assessments of risk of bias and certainty of the evidence were lacking. Further, 42% of abstracts, 58% of plain language summaries, and 67% of summary of findings tables did not mention any harms. We also found that treatment burden and opportunity costs in terms of human resources were rarely considered, while financial costs were more often reported.

We acknowledge that since authors of reviews have many patient-relevant outcomes to consider, it may not be feasible to include all harms we investigated in every review on screening, even for those harms we judged to be relevant to the specific screening intervention investigated. Although the majority of the harms of screening addressed in this paper may be estimated based on the same studies as the benefits, some may require additional searches for other study types or specific methodological considerations when synthesizing the evidence [8], and thus, risks increasing complexity and time to the publication of a review. In other words, the optimal rate might not be 100% for all of the outcomes included in this study. But harms that are important enough to change screening recommendations due to frequency or severity should always be assessed. Further, considering the very low rate of reviews and protocols that addressed even the major harms of screening, we do not believe that the results found in our study can be justified.

### 4.1. Limitations

The most important limitation of this study is that some parts of the analysis required subjective judgment. The judgments we made are available in the Appendix. The numbers and percentages presented should not be interpreted as exact estimates. However, we chose to be conservative in our judgments and are more likely to underestimate the problems of poor reporting of harms than to overestimate them.

For example, we chose to exclude reviews and protocols from the denominator in all cases where overdiagnosis may be a problem but would be difficult or almost impossible to estimate, e.g., due to a coinciding preventive effect of colorectal cancer screening. Further, we chose to consider harms as ‘‘reported’’ even when the methodological rigor was limited. For example, while potential benefits of screening were listed one by one as outcomes of the reviews, harms were commonly gathered under one outcome; ‘‘All adverse effects of screening (for example, psychological harms, overdiagnosis, false positives, and other harms reported in the trials).’’ For most of these harm outcomes, the authors only reported results as mentioned in the trials and did not perform risk of bias assessment, meta-analyses, or an evaluation of the certainty of the evidence. In many of these cases, harms were reported in the result section by a single sentence; ‘‘None of the included trials reported on adverse events.’’ Despite this, we considered all of the harms mentioned as ‘‘reported’’ in the review or protocol.

In two reviews, an increase in incidence due to the intervention was reported, whereas overdiagnosis was not. Since an increase in incidence is generally not the same as overdiagnosis (due to, for example, lead-time bias) [11], we did not count this as reporting on overdiagnosis. In another review, an increase in incidence was included as an outcome, while overdiagnosis was mentioned only ‘‘as reported’’ in one of the included trials, without any risk of bias assessment. In this case, we judged overdiagnosis to be reported in the review, although the methodological rigor was less than for the benefits (for more information, see Appendix).

We want to stress that the results on the number of benefits and harms reported in Abstracts and Plain Language Summaries should be interpreted with caution. This is because what was considered a ‘‘mention’’ of harms and benefits many times included important subjective judgments on our behalf. We chose to be conservative in our judgments and we are more likely to have overestimated the number of harms considered. For example, a statement such as ‘‘none of the included studies reported on adverse events’’ was counted as mentioning one harm.

There is a risk that we missed some reviews of screening interventions since not all of them may have included the word "screening" in the title or as a keyword. We know of one Cochrane review of general health checks [12], which could have been included based on our inclusion criteria but was not picked up by our search strategy.

However, the purpose of this study was not to be exhaustive but to provide an overall understanding of how often harms are reported, and we, therefore, do not believe that this has introduced substantial bias to our results and conclusion [12].

Some of the included reviews were published many years ago, and methods and recognition of harms and their importance in screening may have developed since. However, the reporting of harms was also poor in recently published reviews and protocols. For example, of the 18 protocols and reviews published in 2014 and thereafter, and where overdiagnosis was judged to be relevant, only 3 reported on overdiagnosis, and the corresponding number for overtreatment was 4 of 18 (see Appendix).

### 4.2. Departures from protocol

We planned to investigate the rate of which treatment burden and opportunity costs in terms of financial and human resources were considered in the Discussion or Conclusion section of the reviews. However, we found that the reporting of these aspects was unclear and inconsistent and that it was difficult to judge when these aspects had "been considered" or not. We, therefore, concluded that it was not meaningful to report any numbers for these outcomes. We also planned to compare reporting of harms in reviews published before and after 2013. However, we did not perform such analysis due to small numbers. Further, the majority of reviews were published in the years adjacent to 2013, and since it is unlikely that a few years would make a difference in reporting, any differences picked up between the early and late groups would be difficult to interpret.

### 4.3. Next steps

For patients to be able to decide on whether a medical intervention fits in their life, and for policymakers to be able to make a rational prioritization of community resources, both benefits and harms of a medical intervention need to be assessed, the measurable outcomes need to be estimated, and the certainty of these estimates evaluated [1−3]. For this to be possible, both primary studies and systematic reviews need to address both benefits and harms adequately. The underreporting of the harms of screening we found in this study is problematic both from an individual and an organizational perspective. However, it is not surprising considering that the harms of medical interventions have historically received less attention than the benefits [13]. We believe it is time to change this and that Cochrane can take the lead in this important work.

We call for a broad collaboration to develop reporting guidelines and core outcome sets for primary studies and systematic reviews of screening interventions. To not risk that such guidelines will constitute a bureaucratic barrier to the timely publication of up-to-date evidence, editors and authors of systematic reviews should have an active role in this work to make sure that the practical challenges they will face are addressed. Further, we believe it is of great importance that such reporting guidelines are derived through a broad consensus process with a representation of the public, clinicians, policymakers, methodologists, medical ethicists, as well as people from outside of medicine (for example, social sciences and economics). These should all be free from conflicts of interest in relation to screening interventions.

## 5. Contribution of authors

All authors had access to the data and were responsible for the decision to submit the manuscript. All authors have seen and approved the final text.

## CRediT authorship contribution statement

**Minna Johansson:** Conceptualization, Methodology, Data curation, Formal analysis, Project administration, Supervision, Validation, Visualization, Investigation, Writing - original draft, Writing - review & editing. **Franciszek Borys:** Data curation, Investigation, Writing - review & editing. **Hanna Peterson:** Data curation, Investigation, Writing - review & editing. **Giulia Bilamour:** Data curation, Investigation, Writing - review & editing. **Matteo Bruschettini:** Conceptualization, Methodology, Data curation, Funding acquisition, Investigation, Writing - review & editing. **Karsten Juhl Jørgensen:** Conceptualization, Methodology, Data curation, Formal analysis, Funding acquisition, Investigation, Writing - review & editing.

## Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.09.030.

## References

[1] Harris RP, Sheridan SL, Lewis CL, Barclay C, Vu MB, Kistler CE, et al. The harms of screening: a proposed taxonomy and application to lung cancer screening. JAMA Intern Med 2014;174:281−5.
[2] UK National screening Comittee. Criteria appraising Viability effectiveness appropriateness a Screen programme. Available at www.screening.nhs.uk/criteria. Accessed December 1, 2019.
[3] Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bull World Health Organ 2008;86:317−9.
[4] Heleno B, Thomsen MF, Rodrigues DS, Jorgensen KJ, Brodersen J. Quantification of harms in cancer screening trials: literature review. BMJ 2013;347:5334.
[5] DeFrank JT, Barclay C, Sheridan S, Brewer NT, Gilliam M, Moon AM, et al. The psychological harms of screening: the evidence we have versus the evidence we need. J Gen Intern Med 2015;30:242−8.
[6] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for systematic reviews of interventions

*version 6.0* (updated July 2019). Cochrane 2019. Available at www.training.cochrane.org/handbook. Accessed October 15, 2020.

[7] Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K. Harms are assessed inconsistently and reported inadequately part 1: systematic adverse events. J Clin Epi 2019;113:20−7.

[8] Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K. Harms are assessed inconsistently and reported inadequately part 2: non-systematic adverse events. J Clin Epi 2019;113:11−9.

[9] Johansson M. Harms in Cochrane reviews of screening. PROSPERO 2019. CRD42019138110. Available at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019138110. Accessed October 15, 2020.

[10] Welch HG, Schwartz L, Woloshin S. Overdiagnosed — making people sick in the pursuit of health. Boston: Beacon Press; 2011.

[11] Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol 2007;8:1129−38.

[12] Krogsbøll LT, Jørgensen KJ, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. Cochrane Database Syst Rev 2019;1:CD009009.

[13] Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. PLoS Med 2016;13:e1002127.