



SYSTEMATIC REVIEW

Quality of information in news media reports about the effects of health interventions: Systematic review and meta-analyses [version 1; peer review: 3 approved, 1 approved with reservations]

Matt Oxman ^{1,2}, Lillebeth Larun³, Giordano Pérez Gaxiola ⁴, Dima Alsaied⁵, Anila Qasim⁶, Christopher James Rose ³, Karin Bischoff⁵, Andrew David Oxman¹

¹Centre for Informed Health Choices, Norwegian Institute of Public Health, Oslo, Norway

²Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

³Division for Health Services, Norwegian Institute of Public Health, Oslo, Norway

⁴Cochrane Associated Centre and Evidence-based Medicine Department, Sinaloa's Pediatric Hospital, Culiacan, Mexico

⁵Institute for Evidence in Medicine (for Cochrane Germany Foundation), Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

⁶Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada

V1 First published: 01 Jun 2021, 10:433
<https://doi.org/10.12688/f1000research.52894.1>
 Latest published: 01 Jun 2021, 10:433
<https://doi.org/10.12688/f1000research.52894.1>

Abstract

Background

Many studies have assessed the quality of news reports about the effects of health interventions, but there has been no systematic review of such studies or meta-analysis of their results. We aimed to fill this gap (PROSPERO ID: CRD42018095032).

Methods

We included studies that used at least one explicit, prespecified and generic criterion to assess the quality of news reports in print, broadcast, or online news media, and specified the sampling frame, and the selection criteria and technique. We assessed criteria individually for inclusion in the meta-analyses, excluding inappropriate criteria and criteria with inadequately reported results. We mapped and grouped criteria to facilitate evidence synthesis. Where possible, we extracted the proportion of news reports meeting the included criterion. We performed meta-analyses using a random effects model to estimate such proportions for individual criteria and some criteria groups, and to characterise heterogeneity across studies.

Results

We included 44 primary studies in the qualitative summary, and 18 studies and 108 quality criteria in the meta-analyses. Many news reports gave an unbalanced and oversimplified picture of the

Open Peer Review

Reviewer Status

	Invited Reviewers			
	1	2	3	4
version 1				
01 Jun 2021	report	report	report	report

1. **Raymond Moynihan**, Bond University, Gold Coast, Australia
2. **Joel R. Lexchin** , York University, Toronto, Canada
 University of Toronto, Toronto, Canada
 University Health Network, Toronto, Canada
3. **Amanda Wilson** , University of Technology Sydney, Sydney, Australia
4. **David Henry** , Bond University, Gold Coast, Australia

potential consequences of interventions. A limited number mention or adequately address conflicts of interest (22%; 95% CI 7%-49%) (low certainty), alternative interventions (36%; 95% CI 26%-47%) (moderate certainty), potential harms (40%; 95% CI 23%-61%) (low certainty), or costs (18%; 95% CI 12%-28%) (moderate certainty), or quantify effects (53%; 95% CI 36%-69%) (low certainty) or report absolute effects (17%; 95% CI 4%-49%) (low certainty).

Discussion

There is room for improving health news, but it is logically more important to improve the public's ability to critically appraise health information and make judgements for themselves.

Keywords

news, news media, news reports, health news, systematic review, meta-analysis, infodemic

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Matt Oxman (matt@mattoxman.com)

Author roles: **Oxman M:** Conceptualization, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation; **Larun L:** Investigation, Writing – Review & Editing; **Pérez Gaxiola G:** Investigation, Writing – Review & Editing; **Alsaid D:** Investigation, Writing – Review & Editing; **Qasim A:** Investigation, Writing – Review & Editing; **Rose CJ:** Formal Analysis, Writing – Review & Editing; **Bischoff K:** Investigation, Writing – Review & Editing; **Oxman AD:** Conceptualization, Investigation, Methodology, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Oxman M *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Oxman M, Larun L, Pérez Gaxiola G *et al.* **Quality of information in news media reports about the effects of health interventions: Systematic review and meta-analyses [version 1; peer review: 3 approved, 1 approved with reservations]** F1000Research 2021, 10:433 <https://doi.org/10.12688/f1000research.52894.1>

First published: 01 Jun 2021, 10:433 <https://doi.org/10.12688/f1000research.52894.1>

Introduction

For decades, researchers have studied and criticised the quality of health information in the mass media. In 1972, *The New England Journal of Medicine*—the “most prestigious” general medical journal in the world¹—published an “examination” of health information on American television (TV).² The paper ended in reproach: “The potential of television to inform, instruct and educate should not continue to be wasted and abused, but rather should be used as an integral part of a plan designed to deliver better health care.”

Misleading health information in the mass media continues to be a hot topic. In 2020, the World Health Organization reported that the Coronavirus disease (COVID-19) had been accompanied by an “infodemic”: “an over-abundance of information—some accurate and some not—that makes it hard for people to find trustworthy sources and reliable guidance when they need it.”³

Many studies of health news have focused on “traditional” news media, including print, broadcast, and news websites. With the advent of social media, traditional news media have increased competition as a source of lay health information. However, people still get health information from the latter, directly or via social media. Chew and Eysenbach found that about 60% of more than 5000 tweets about the H1N1 outbreak in 2009 included links and about a quarter of those links were to news websites, the most common destination.⁴ In comparison, less than 5% of links were to government and public health agencies or to other social media, combined.

Many of the studies focusing on news media have further focused on news reports about the effects of health interventions. This includes “modern” medicine (aka. “academic”, “conventional” or “Western” medicine); “alternative” medicine (aka. “complementary”, “traditional” or “natural” medicine); screening; surgery; devices; diet; exercise; lifestyle interventions; and health systems and policies. We were unable to find a systematic review of such studies (see extended data - S1, S2 and S3 Files¹⁰⁴).⁵ We then aimed to fill this gap, providing meta-analytical estimates of the quality of news reports about the effects of health interventions, and potentially informing further and related research.

The first part of this study was submitted by the first author, MO, to the University of Oxford as his dissertation, in partial fulfilment of the requirement for the award of Master of Science (MSc) in Evidence-Based Health Care, under the title “Criteria used to measure the quality of news media report about the effects of health interventions: Systematic review” (see extended data - S4 File¹⁰⁴).

Objectives

Primary:

- To assess the quality of information in news media reports about the effects of health interventions.

Secondary:

- To assess, map and group criteria used to measure the quality of information in news media reports about the effects of health interventions.

Methods

This review was registered in the International Prospective Register of Systematic Reviews (**PROSPERO**) (ID: [CRD42018095032](https://doi.org/10.1186/1745-6215-42018095032)). The protocol was published on the **Informed Health Choices project** website as “Quality of news media reports about the effects and costs of health interventions: Systematic review protocol” (see extended data - S1, S2 and S3 Files¹⁰⁴).⁵ Appendixes to the protocol were published in a separate compressed folder, on the same website. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (**PRISMA**) Checklist⁶ has been included in the *Reporting guidelines* for this review.

Eligibility criteria

Primary studies had to satisfy four eligibility criteria to be included in the review (**Table 1**): two related to the sample of news reports included in the study; one related to the outcome measure; and one related to reporting of the study. We included studies in any language in the qualitative summary, but we only included studies in English or Norwegian in the meta-analyses (the two languages spoken fluently by the primary investigator, MO). We did not place any explicit limitation on study design, but—in effect—case studies of a single news report and purely qualitative studies were ineligible.

The first eligibility criterion was that the sample include content labelled by the researchers as “news”, taken from newspapers or magazines (print); television, radio or podcasts (broadcast); or dedicated news websites (online). We placed no limits on where in the world or when the news reports were published.

Table 1. Eligibility criteria for primary studies.

Sample	<i>Eligibility Criterion 1:</i>	Includes content labelled by the researchers as “news”, taken from newspapers or magazines; television, radio or podcasts; or dedicated news websites
	<i>Eligibility Criterion 2:</i>	Includes news reports about the effects of health interventions
Outcome measure	<i>Eligibility Criterion 3:</i>	Includes at least one explicit, prespecified and generic criterion for quality
Reporting	<i>Eligibility Criterion 4:</i>	Specifies the sampling frame, and selection criteria and technique

Second, the sample had to include news reports about the effects of health interventions. Other types of health news include non-reports (e.g. features or opinion pieces) and other categories of news reports about health (e.g. reports about the health effects of exposures or about health-related ethical or legal issues). We defined “health intervention” as any action intended to improve the health of individuals or communities, as per the Glossary of Evaluation Terms for Informed Treatment choices (GET-IT).⁷ We defined “effect” as a negative or positive change or difference in a health outcome, again as per GET-IT.

Third, the researchers had to have used at least one explicit, prespecified and generic criterion to measure quality, herein called a quality criterion. “Quality” was defined as how conducive the report is to informed decisions or, conversely, how misleading it is. Information about the effects of health interventions can be directly misleading—such as claims that an intervention causes an outcome when it is only associated with it—or it can be misleading by omission, such as omitting absolute effect estimates, particularly when the baseline risk is small.

“Criterion” was defined as a standard of quality that could be satisfied or not. “Generic” meant that the quality criterion was not specific to an intervention (e.g. a drug), category of interventions (e.g. “modern” medicine), or condition (e.g. cancer), or specific to any evidence (e.g. findings from a trial). We excluded studies if they only measured the quality of the evidence cited in the news reports; only measured the factual accuracy of the reports; or only described the reports, as opposed to explicitly assessing their quality. We included quality criteria related to cost of the intervention.

The fourth and final eligibility criterion was that the study specified the sampling frame (where the news reports were sampled from), the selection criteria for the news reports, and the selection technique (how the reports were sampled).

To clarify the eligibility criteria, three reviewers piloted them on a sample of five studies from a list of potentially eligible studies of which we were aware before finalising the protocol (Appendix 1 of the protocol; extended data - S2 File¹⁰⁴).⁵

Search strategy

We searched for eligible primary studies indexed in PubMed on May 24, 2018, and Google Scholar on June 21 and 22, 2018. We searched directly in Open Grey and Grey Literature Report on June 22, 2018, and in ProQuest Dissertations & Theses (Global Full text plus UK and Ireland abstracts) on June 25, 2018. We updated the PubMed search on Aug. 30, 2019.

We conducted citation searches for papers describing the development of tools for assessing the quality of news reports about health: the Index of Scientific Quality;⁸ the Quality Index for health-related Media Reports;⁹ and a checklist for improving drug information in the general press.¹⁰ We also checked the reference lists of those three papers. Finally, we conducted citation searches and reference lists checks for an arbitrary selection of eligible studies. S5 File in the extended data¹⁰⁴ is a detailed description of the search strategy, including search strings.

Study selection

Screening search results for inclusion in the qualitative summary consisted of two rounds. First, two reviewers screened titles and abstracts. Second, two reviewers screened full texts of studies included after the first round. The two discussed and resolved disagreements after each round. Where two reviewers could not reach a consensus, a third was brought in to arbitrate. Screening reference lists and the results of citation searches included a third, initial round, wherein only one reviewer screened titles. We also screened all of the potentially eligible studies of which we were aware before finalising the protocol (Appendix 1 of the protocol; extended data – S2 file¹⁰⁴). We screened search results using Rayyan.¹¹

If a study included a mix of news reports about the effects of health interventions and other types of health news, we excluded the study from the meta-analyses unless results for the news report about the effects of health interventions were

reported separately for at least one eligible quality criterion. Further, we excluded studies from the meta-analyses if they did not include any eligible quality criteria or any such criteria for which results were adequately reported (see the section “Assessing criteria”). Judgements about excluding studies from the meta-analyses were made independently by two reviewers, who then discussed and resolved disagreements.

Data extraction

Data were extracted by a single reviewer and entered into spreadsheets (Microsoft Excel, Version 16.49). We only extracted reported data. For example, if a study did not include the country or countries in which the sample of news reports were published, we did not attempt to code the individual news reports by country or contact the study authors for that information. A second reviewer checked a sample of the extracted descriptive data and checked all the extracted quality criteria that were included in the meta-analyses, as well as results for those criteria.

For the qualitative summary, we extracted descriptive data about the samples of news reports and the tools (sets of quality criteria) used to assess the reports. Data about the sample of news reports included: the categories of media in which the reports were published; the countries in which they were published; the time period in which they were published; the categories of health interventions that were the subject of the reports; and the financial models of the news outlets that published the reports. We also extracted study objectives, sampling frame, selection criteria and technique, and reported subgroup analyses.

All categories and the time periods were predetermined in the protocol. The categories of media were newspaper; magazine; radio; podcast; television; and news website. The categories of intervention were: “modern” medicine; “alternative” medicine; screening; surgery; devices; diet; exercise; lifestyle; and systems and policies. The two categories of financial model were non-commercial (public or independent) and commercial.

We extracted quality criteria verbatim. For the meta-analyses, we extracted all response options for each included quality criterion, the number of relevant news reports (reports about the effects of health intervention) to which the quality criterion was applied, and the results for those reports. Some quality criteria were not applied by researchers to the total sample of relevant reports in a given study, meaning the researchers did not consider the quality criterion universally applicable.

Again, if a study included a mix of news reports about the effects of health interventions and other types of health news, we only included results for the former. We excluded studies if those results were not reported separately.

For the meta-analyses, we were interested in the proportion of relevant news reports that satisfied a given quality criterion. If this was not reported explicitly, we imputed the proportion using reported data, if possible. We excluded quality criteria if the proportion was not explicitly reported and it was impossible to calculate the proportion based on the reported data—for example, if only mean scores were reported.

Assessing criteria

After selecting studies to be included in the meta-analyses (Figure 1), two authors assessed criteria used in those studies: both quality criteria and other criteria, such as descriptive criteria or criteria measuring factual accuracy, using spreadsheets. We excluded inappropriate criteria, as well as criteria for which it was not possible to extract results for only news reports about the effects of health interventions.

We assessed each criterion in each study, even if the researchers used a tool used in a previous study, in case it had been modified. For example, the Media Doctor Australia tool is used in several studies, including those by Smith *et al.*¹² and Bonevski *et al.*¹³ However, while Smith *et al.* asked whether there is “objective evidence to support the treatment”, Bonevski *et al.* asked whether the study methodology is reported, which is more specific.

Two reviewers independently assessed criteria, then discussed and resolved any differences. We excluded criteria if they were non-generic or if they assessed factual accuracy, if they were descriptive, or if they were not clear or sensible (i.e. if their face validity was inadequate¹⁴). We also excluded global scores and composite quality criteria because they can be difficult to interpret, and they are difficult to compare across studies using different scoring systems.

Mapping and grouping quality criteria

We mapped and grouped quality criteria, using spreadsheets, to be able to synthesise results across studies using different tools. For the mapping, we used the Informed Health Choices (IHC) Key Concepts framework. The most recent version of the framework was published on IHC website in 2019 and includes nine high-level concepts and 49 low-level concepts

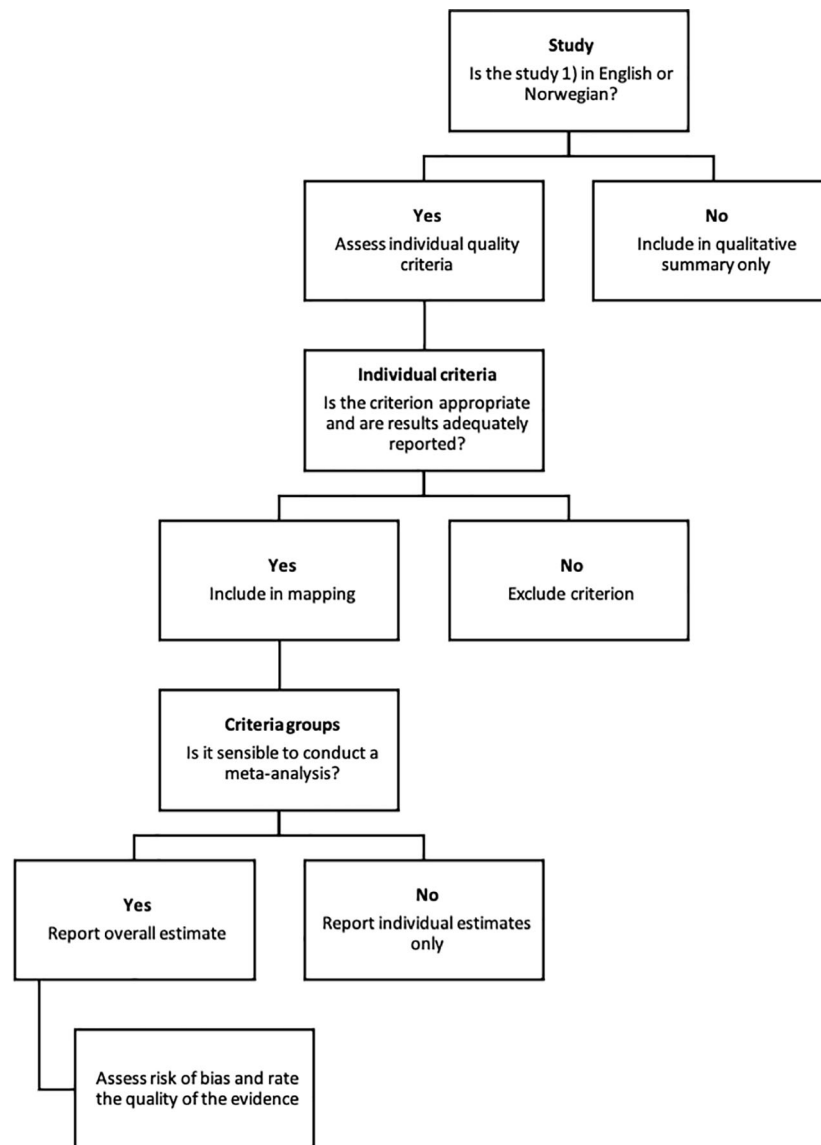


Figure 1. Post-study selection process.

organised in three overarching groups (see extended data - S6 File¹⁰⁴).¹⁵ The IHC Key Concepts are essential concepts for assessing information about the effects of health interventions and making informed choices.¹⁵⁻¹⁷ The framework is based on a mix of evidence and logic. It has a combination of characteristics that make it unique, as well as uniquely useful for this study. First, it has been developed systematically, transparently and iteratively. Second, it only includes concepts that are intended to be relevant to the general public (typically the target audience of health news). And third, it is intended for mapping and assessing measurement tools (amongst other purposes). Moreover, the starting point for the framework included checklists for journalists and input from journalists.¹⁸

Two reviewers independently mapped the included quality criteria against the low-level IHC Key Concepts, then discussed and resolved any disagreements. Finally, one reviewer grouped the criteria, informed by the results of the mapping, and a second reviewer assessed whether the groups were sensible (i.e. face validity¹⁴).

Meta-analysis

Where an included quality criterion had more than two response options, one reviewer dichotomised the categories and a second reviewer assessed the dichotomisation. For example, for studies that used the Index of Scientific Quality, where response options range from a score of one to five, we considered a score of four or five as satisfying the quality criterion.

Furthermore, we reframed negative quality criteria as positive, and paraphrased all quality criteria, for the data to be consistently framed and worded, and for the criteria to be succinctly labelled. For example, Bubela *et al.* report the number of news reports that “did not” specify whether a study was randomized, which was reframed as the number of news reports that did specify whether a study was randomized.¹⁹

For each individual quality criterion and group of criteria, we estimated mean proportions with 95% confidence intervals (CIs), using a random effects model, quantifying precision using 95% confidence intervals. Where criteria groups included more than one criterion from the same study, we included results for both criteria. We performed statistical analysis using Stata 16 (StataCorp LLC, College Station, Texas, USA). Each of the included studies provided data on the number of relevant news reports assessed and how many of those reports met a given criterion. For each study and quality criterion for which necessary data were available, we estimated the proportion of relevant news reports meeting the criterion. These point estimates and 95% confidence intervals were transformed to the logit scale for meta-analysis (Stata's default scale for meta-analysis of proportions). We had anticipated substantial between-study heterogeneity and therefore used random effects meta-analyses. We estimated the percentage of variance that can be attributed to between-study heterogeneity rather than chance (I^2 , expressed as a percentage) and calculated P-values to test the null hypotheses of no between-study heterogeneity using Cochran's Q statistic. We back-transformed meta-analytical estimates from the logit scale to present estimates as percentages (in forest plots summarizing meta-analytical results from different criteria groups) or proportions (in forest plots for individual groups).

Two reviewers assessed whether what the different quality criteria in each group measured was similar enough that it was sensible to report an overall estimate. If a group included a single quality criterion, we did not report an overall estimate.

Risk of bias

Two reviewers independently assessed the risk of detection bias for each criterion included in the meta-analyses, then compared judgements and resolved disagreements. Detection bias was assessed using three questions:

1. Was the criterion applied by at least two researchers independently?
2. Were the researchers blinded in terms of the journalist and news outlet?
3. Does the criterion require substantial judgement?

If the authors did not state how many researchers applied the criterion or whether the researcher(s) did so independently, we recorded the answer to the first question as “unclear” and counted it as a “no”. If from looking at the criterion in and of itself, it was unclear how much judgement was required, we looked at whether the researchers reported or referenced any guidance for applying it. If they did, we looked at how detailed and specific the guidance was. If the researcher(s) applying the criterion received training, but the authors did not report what this entailed—i.e. what guidance was provided—we considered the criterion to require substantial judgement.

We considered the risk of detection bias as low if the criterion did not require substantial judgement. We considered the risk of detection bias as moderate if the criterion required substantial judgement, but at least two researchers applied the criterion independently and they were blinded in terms of the journalist and outlet. We considered the risk of detection bias as high if the criterion required substantial judgement and only one researcher applied it or the researchers were not blinded.

If there was a high risk of bias for criteria providing $\geq 50\%$ of the weight in a given criterion group, we considered there to be a high risk of bias for the overall estimate.

Rating the quality of the evidence

We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach as a framework for rating the quality of evidence.²⁰ GRADE considers five factors that can lower the certainty of evidence for effect estimates: risk of bias, inconsistency of results, indirectness of evidence, imprecision, and publication bias. We did not consider publication bias, as we are not aware of any research documenting publication bias in studies of the quality of news reports or similar studies. We also did not consider directness, which was irrelevant.

Again, if there was a high risk of bias for criteria providing $\geq 50\%$ of the weight in a given group, we considered there to be a high risk of bias for the overall estimate. If a CI is wider than 0.25 (a quartile), we considered there to be important

imprecision, based on guidance from the Cochrane Effective Practice and Organization of Care (EPOC) group.²¹ We used an I^2 of $>75\%$ as a rule of thumb for there being substantial inconsistency.

GRADE factors for increasing the certainty of the evidence were irrelevant.

Deviations from the protocol

We report the main deviations from the protocol here. In S7 File (see extended data¹⁰⁴), we provide detailed information about all deviations. Some of the deviations were in part due to the originality of our research question in terms of a systematic review and meta-analysis, meaning there was lack of methodological guidance or precedence for the study as a whole. In other words, we were in “unchartered waters”. Other deviations were due to limited time and resources. This review had no funding and as noted in the background section, the first part of the review was MO’s MSc dissertation, meaning he was required to do as much as possible of that work himself.

In terms of search strategy, we did not search **Scopus** as planned, aside from citation searches, since the time saved outweighed the likelihood of identifying additional eligible studies, based on the results of searching the other databases. Also, we only conducted reference list checks and citation searches for arbitrary samples of included studies, due to limited time and resources, as well as decreasing yield (extended data – S5 File¹⁰⁴).

In terms of eligibility, we made explicit that at least one of the criteria that researchers in a given study used to assess quality had to be prespecified and generic, and used specifically to assess the quality of news reports, not describe their content or assess the quality of underlying evidence. As mentioned, we included studies in any language in the qualitative summary, but only included studies in English and Norwegian (the two languages spoken fluently by the primary investigator, MO) in the meta-analyses.

In terms of data extraction, a second reviewer checked all data extracted for the meta-analyses, but only a sample of data extracted for the qualitative summary due to limited time and resources. There were no errors in the sample. In terms of assessing risk of bias, we were originally going to assess selection bias by looking at how news reports were sampled: randomly, sequentially or other. However, we found this told us more about applicability than bias, since samples typically consisted of all news reports that met the researchers’ eligibility criteria, rather than a sample of those reports. Therefore, we only assessed the risk of detection bias, as described earlier.

In terms of synthesising results and rating the quality of the evidence, we have not reported overall estimates for groups of quality criteria if what those criteria measure is too different for a synthesis to be sensible, despite the criteria being clearly related. Nor have we reported overall estimates for groups that include a single criterion. Further, we did not rate the quality of the evidence in these cases. Where we do report an overall estimate and rate the certainty of the evidence, if we assessed the risk of bias as high for criteria providing $\geq 50\%$ of the weight in a given group, we assessed the risk of bias as high for the group as a whole. We did not consider assessing risk of bias for whole criteria groups in the protocol.

Results

We screened 2063 unique records and retrieved and assessed 140 full texts for eligibility, as shown in the PRISMA flow diagram (Figure 2). We identified 44 primary studies for inclusion in the qualitative summary (see extended data - S1 Table¹⁰⁴): 31 in English, five in German, four in Spanish, three in Norwegian and one in Slovenian. We excluded most (78) full-text articles from the qualitative summary based on one or more of the four eligibility criteria. We excluded another 16 full-text articles that were not research (e.g. commentaries), and two duplicates: one doctoral thesis that is a composite of other studies that we assessed²² and one working paper²³ superseded by the final publication.²⁴ The third eligibility criterion—that the researcher had used at least one explicit, prespecified and generic quality criterion—was least often satisfied. We excluded 47 full-text articles solely for failing to satisfy that eligibility criterion.

Of the 44 studies included in the qualitative summary, we included 18 (41%) in the meta-analyses. Of the 26 studies only included in the qualitative summary, we excluded nine (20%) from the meta-analyses for being in a language other than English or Norwegian; 16 for inadequate reporting (36%); and one (2%) where the sample of news reports²⁵ was included in a larger sample in a later study.²⁶ Only one of the 18 studies included in the meta-analyses was in Norwegian,²⁷ meaning the rest were in English.

Study characteristics

Samples

Table 2 is an overview of the samples of news reports in the studies included in the qualitative summary. Descriptive data for each study can be found in S2 Table in the extended data.

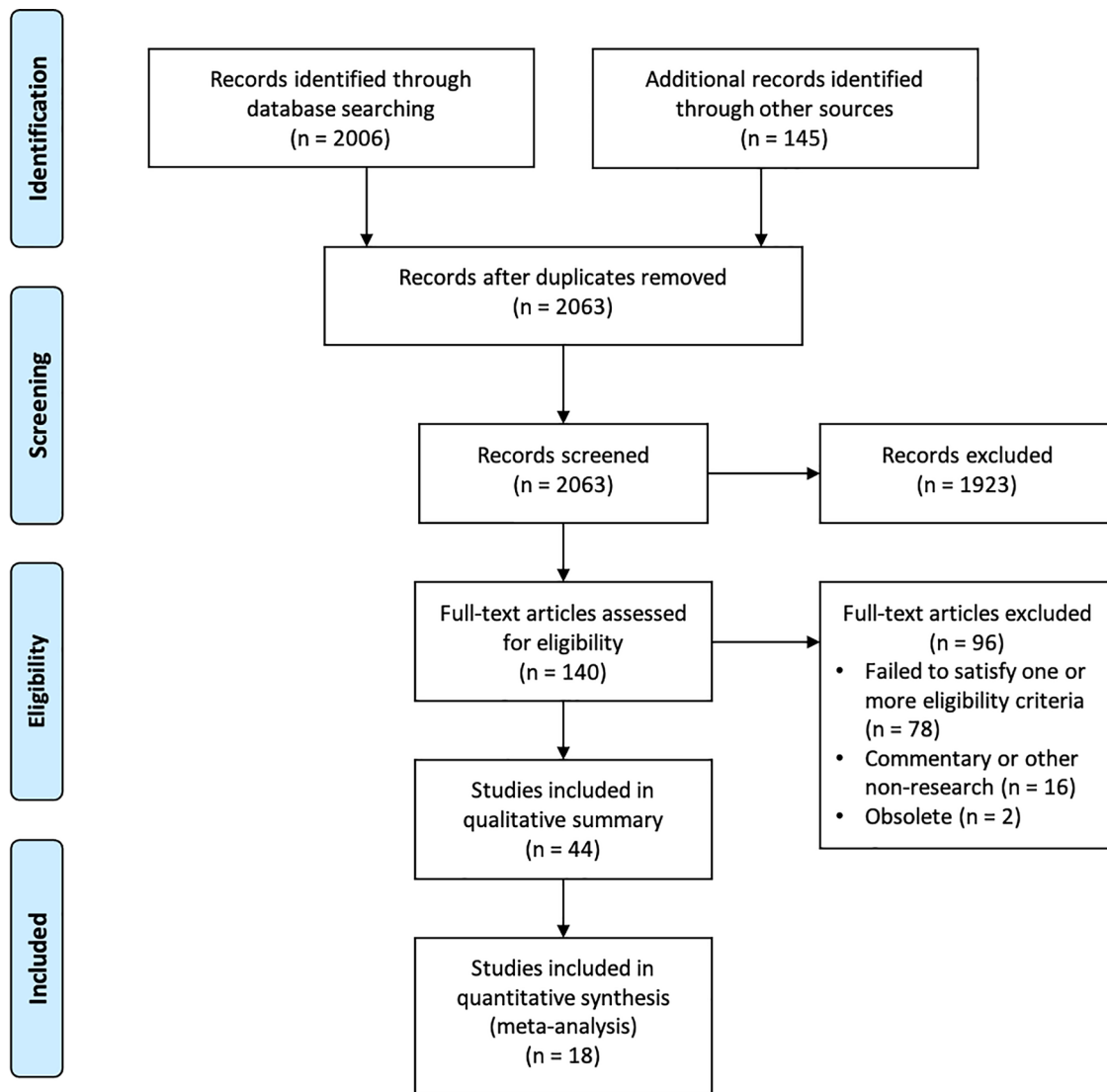


Figure 2. PRISMA flow diagram.

Sample sizes varied from 17 to 3096 units. We use “units” here since some samples include content besides news reports, such as non-news websites,²⁸ or include health news that is not news reports about the effects of health interventions, for example news features²⁹ or news reports about the effects of exposures, such as recreational cannabis (as opposed to medical cannabis).³⁰

Newspaper was the most common medium, with newspaper content explicitly included in 35 of the 39 studies (90%) in which a medium was specified (Table 2). However, the number of studies that included newspaper content versus content from news websites should be interpreted with caution, since newspaper content may also have been published online. In the study with the second-largest sample, Walsh-Childers *et al.* write that news reports “must have been published on the news organization’s website,” but categorise those reports only as newspaper content.²⁶

The United States (US) was the most common country in terms of where content assessed in the included studies was published. Of the 40 studies where at least one country was specified, 12 samples (30%) included content published in the US (Table 2). Twenty-five (63%) of those 40 studies included content published in English-speaking countries (see extended data - S2 Table¹⁰⁴).

We categorised time periods according to what authors reported. This was in some cases the period in which news reports were published and in others the period in which reports were assessed. Differences between the two are likely minor,

Table 2. Characteristics of samples included in the qualitative summary.

Medium ^a	Studies ^b	Country	Studies ^b	Time period ^c	Studies ^b	Intervention category ^d	Studies ^b
Magazines	10	Argentina	1	1990-1999	7	“Alternative” medicine	4
News websites	8	Australia	5	2000-2009	23	Devices	3
Newspapers	35	Canada	5	2010-2019	19	Diet	7
Other (Print) ^e	1	Germany	5			Exercise	4
Podcasts	0	Italy	2			Lifestyle	4
Radio	1	New Zealand	2			“Modern” medicine	24
TV	7	Norway	2			Other	7
Unspecified	5	Slovenia	1			Screening	5
		South Africa	1			Surgery	6
		Spain	3			Systems and policies	3
		Trinidad and Tobago	1			Unspecified	12
		United Kingdom	7				
		United States	12				
		Vietnam	1				
		Unspecified	4				

^aAs per categories pre-specified in the protocol: newspapers, magazines, TV, radio, podcasts, and news websites.

^bThe same study may be counted in several categories.

^cTime period when news reports were published or studied, depending on what authors reported.

^dAs per categories pre-specified in the protocol: “modern” medicine, “alternative” medicine, screening, surgery, devices, diet, exercise, lifestyle, and systems and policies.

^eThe authors do not specify whether the news reports were sampled from newspapers, magazines or both.

since all studies focused on the recent past (relative to the time that the study was conducted). The earliest year of publication was 1996.³¹

“Modern” medicine was included in 24 of 32 samples (75%) where a category of interventions is specified. In some cases, it was not possible to use the categories pre-specified in the protocol for this review. For example, Haneef *et al.* included news reports about “non-pharmacological” interventions,³² which could include “Alternative” medicine, devices, diet, etc. In these cases, we used the category “Other”.

Financial model was not specified in any of the included studies—i.e. it was not explicitly reported whether the news outlets that published news reports included in the sample were commercial or non-commercial (public or independent).

Tools

The tools used to assess the quality of news reports in each study are presented with the descriptive data in S2 Table in the extended data.¹⁰⁴ Most studies used a tool that was, to varying degrees, original or modified. Only one tool was used in >4 of the 44 studies included in the qualitative summary: the Index of Scientific Quality, which was used in seven studies (16%).^{28,30,31,33–36} However, even more studies used tools that were to some extent, directly or indirectly, based on the criteria developed by Moynihan *et al.*³⁷ This includes the four studies that used the Media Doctor Australia tool^{12,13,38,39} and the 3 studies that used the Health News Review tool,^{25,26,40} which was based on the Media Doctor Australia tool.²⁵ Two studies used two separate tools.^{38,41} We did not identify any studies that used the Quality Index for health-related Media Reports⁹ or the checklist for improving drug information in the general press.¹⁰

Objectives, sampling frames, selection criteria and technique, and reported subgroup analyses

The explicit or implicit objectives of studies included in the qualitative summary are presented in S3 Table, the sampling frames and selection criteria and techniques in S4 Table, and the variables used in the reported subgroup analyses in S5 Table (see extended data¹⁰⁴). Detail and clarity varied. For example, some studies reported explicit inclusion and exclusion criteria for news reports, while others simply stated that they assessed news reports that included health or medical information.

Thirty-one studies (70%) reported at least one subgroup analysis, exploring differences in quality between subgroups of news reports. The most common overarching variables were ones related to time, medium, and outlet. How subgroups were categorised for the same overarching variable, as well as what categories were included, varied. For example, some studies included subgroup analyses based on specific intervention (e.g., comparing the quality news reports about one medicine with news reports about another medicine). Others included subgroup analyses based on the category of intervention (e.g., comparing news reports about “modern” medicines versus news reports about other categories of interventions). Furthermore, the specific interventions or categories of interventions varied.

Assessment, mapping and grouping of criteria

Two reviewers assessed all 208 criteria applied to news reports in the 18 studies included in the meta-analyses. Not all criteria were measures of quality. Of the 208 criteria, 77 (37%) were excluded: 66 (32%) for being non-generic or descriptive, for measuring factual accuracy, or for not being clear or sensible (having inadequate face validity); and 11 (5%) for being global scores or composites. We provide examples in Table 3. Quality criteria were judged not to be sensible if they were in conflict with the IHC Key Concept and research evidence, for example a criterion assessing whether research presented in a news report is peer-reviewed.^{40,42} The same two reviewers mapped the remaining 131 of 208 quality criteria (63%) against the IHC Key Concepts. Based on the results of the mapping, we established 19 groups of quality criteria (criteria groups), as well as eight subgroups.

In S6 Table in the extended data,¹⁰⁴ we present the number of quality criteria included in the meta-analyses per relevant IHC Key Concept. We were unable to identify or include quality criteria related to several IHC Key Concepts that are important for communicating information about the effects of health interventions in the news media or otherwise. For example, there were no criteria related to the concept “Average differences between treatments can be misleading” (2.3c). In contrast, a checklist for people communicating evidence-based information about the effects of healthcare interventions includes the item “Help your audience to avoid misinterpreting continuous outcome measures,” granted the checklist is not for journalists.⁴³

Table 3. Examples of criteria individually excluded from meta-analyses.

Reason	Criterion	Explanation	Citation
<i>Non-generic</i>	“Potential risks or harm”; “Accordance with available evidence”	The criterion measures whether risks or harms mentioned in the news report are consistent with specific evidence, not whether risks or harms are mentioned or discussed.	Robinson 2018
<i>Descriptive</i>	“Picture/Illustration”	The criterion is a descriptive variable about whether the news report includes a picture or illustration, not intended to be a measure of quality.	Finer 1999
<i>Measuring factual accuracy</i>	«Is the treatment genuinely new?»	The criterion is a measure of whether information in the news report is true, as opposed to informative (assuming it is true).	Bonevski 2008
<i>Unclear</i>	«Inadequate claim of safety»	There is no explanation of “inadequate”, making it impossible to assess whether the criterion is sensible.	Yavchitz 2012
<i>Not sensible</i>	«Was the research peer reviewed?»	Peer review is a poor indicator of quality. ⁴²	Stassen 2016
<i>Global score or composite</i>	«News with at least one misleading reporting»	This was a composite of several criteria measuring different types of misleading news reporting.	Haneef 2015

Synthesis of results

Of the 131 quality criteria included in the 19 groups, we excluded 22 (17%) from the meta-analyses for inadequate reporting of results, meaning it was impossible to impute the proportion of relevant news reports that satisfied the criterion based on the reported data. The distribution of the remaining 108 criteria amongst the 19 main groups and eight subgroups, is presented in Table 4, with the order of the groups based on the IHC Key Concept framework. In one case, we included a criterion in two groups: the third criterion in the study by Stassen,⁴⁰ which we included in Criteria related to qualitative descriptions of effects and Criteria related to pros and cons. Several criteria from the same study were included in three of the 19 main groups (16%) and four of the eight subgroups (50%). Criteria related to “disease mongering”¹² stem from a paper by Moynihan *et al.*⁴⁴

The primary meta-analyses are summarised in Figure 3, excluding subgroups. The samples of news reports in two of the studies included in the meta-analyses overlap.^{12,13} For groups that included a criterion from both of these studies, we performed primary analyses that included only the largest and most-recent study by Bonevski *et al.*,¹³ and an ad hoc sensitivity analysis that also included results from the study by Smith *et al.*¹² The sensitivity analyses are summarised in Figure 4, again excluding subgroups. There were no important differences between the overall estimates and CIs for the primary analyses versus the sensitivity analyses. Forest plots for each criteria group, including individual estimates and CIs for each criterion, are presented in S7 File in the extended data.¹⁰⁴

For most of the main groups (14 of 19) and a minority of the subgroups (two of eight), we did not consider it sensible to report overall estimates because—although the criteria in the group are clearly related—what they each measure was too different for it to be sensible to synthesise results. For example, in the group Criteria related to associations and

Table 4. Criteria groups. N = criteria included in the meta-analyses.

Criteria related to	N	Subgroups	N
Associations and randomisation	3		
Need for comparisons	3		
Consistency between studies	1		
“Disease mongering” ⁴⁴	5		
Conflicts of interest	7		
Opinions	1		
Study design and risk of bias in general	5		
Blinding	1		
Loss to follow-up	1		
Selective reporting	2		
Qualitative descriptions of effects ^a	8	Quantification of effects ^a	6
		Misleading language	2
Absolute effects	5		
Play of chance	4		
Subgroup analyses	1		
Lack of evidence	1		
Options	7		
Extrapolations ^a	21	Extrapolation (outcome)	2
		Extrapolation (population) ^a	8
		Extrapolation (intervention)	8
Animal studies	1		
Pros and cons ^a	32	Harms ^a	14
		Cost ^a	11
		Benefits	2

^aIncludes multiple criteria from the same study.

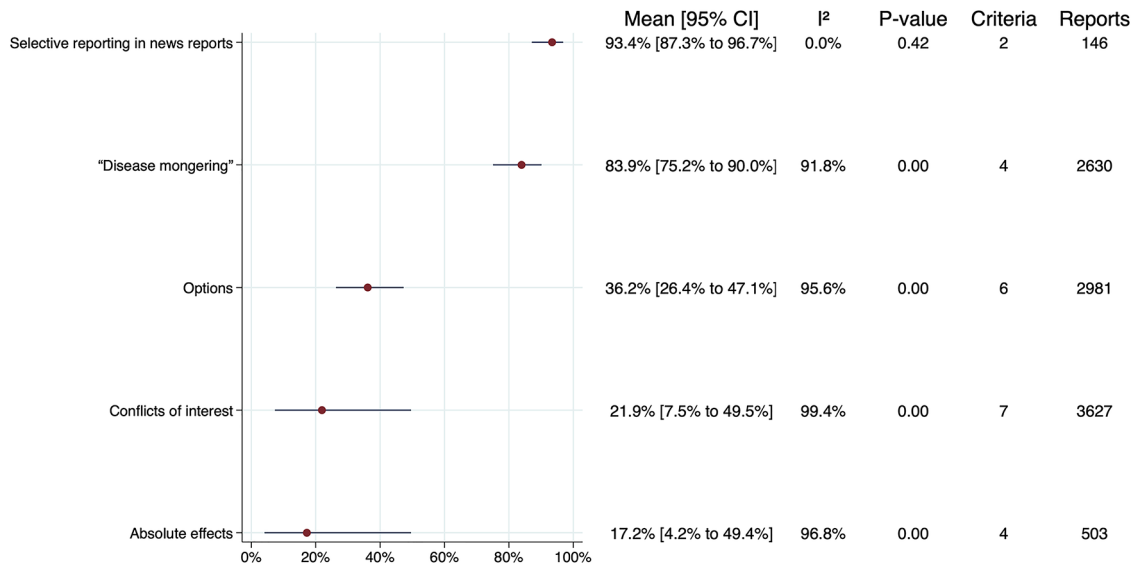


Figure 3. Summary of primary meta-analyses.

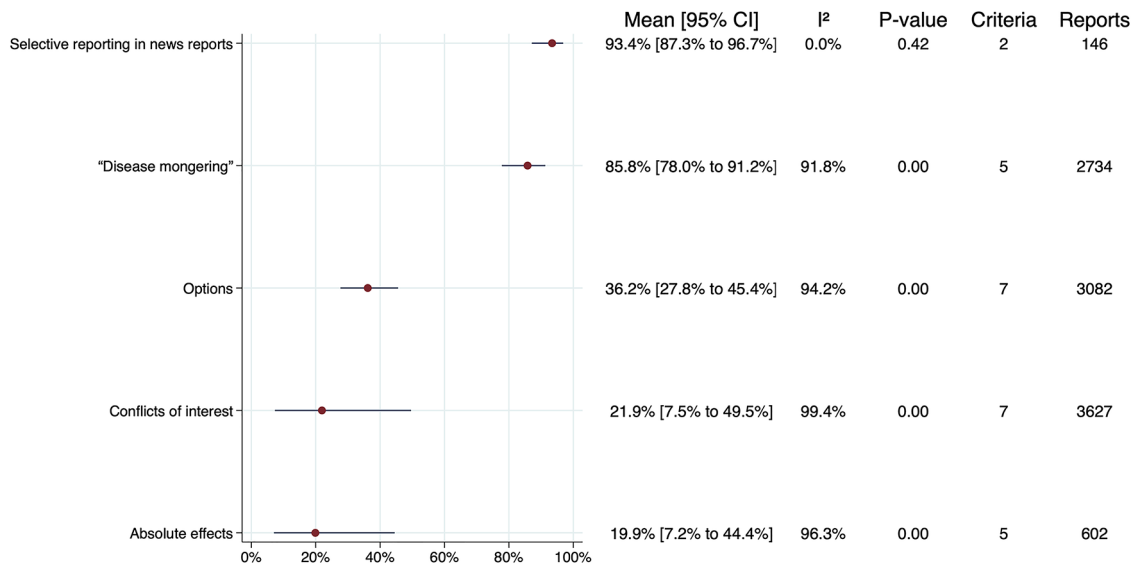


Figure 4. Summary of sensitivity analyses.

randomization, the ninth criterion from the study by Bubela *et al.* measured whether the news report specifies that a trial was randomized or not, whereas the 12th criterion from the study by Haneef *et al.* measures whether there is a causal claim based on a non-randomized study. The two are clearly related, but at the same time, synthesising results would clearly not be sensible. Leaving out whether a study was randomized might be misleading by omission, but this can be compensated for by distinguishing causation and association and pointing out potential confounding. On the other hand, making a causal claim based on a non-randomized study is directly misleading. Further, we do not report overall estimates for the groups that include a single criterion.

Due to heterogeneity, as well as inadequate reporting and small sample sizes in some studies, we did not conduct any of the six subgroup analyses considered in the protocol, exploring differences in quality between news reports:

1. in different media (broadcast vs. other),
2. from different decades,

3. by commercial vs. non-commercial outlets,
4. by health or science journalists vs. others,
5. in low-income countries vs. middle or high-income countries, or
6. in broadsheet vs tabloid newspapers.⁵

Risk of bias

The individual risk of bias assessments for each of the 108 criteria included in the meta-analyses are presented in S9 File in the extended data.¹⁰⁴ We considered there to be a high risk of bias in 13 cases (11%), including in the case of the one criterion included in two criteria groups (the third criterion in the study by Stassen⁴⁰). The 13 criteria for which we considered there to be a high risk of bias are the 13 that required substantial judgement to be applied, according to our assessment. In 12 of those 13 cases, it was—in addition—unclear whether researchers were blinded to the publication and journalist behind the news report. In the remaining case, it was clear that they were not blinded. In four of the 13 cases, it was also unclear whether the criterion was applied independently by at least two researchers.

Quality of the evidence

We only assessed the quality of the evidence for the five main groups and six subgroups where we report an overall estimate. Quality varied from low to high. In all but two of 10 cases, we downgraded quality for substantial inconsistency. In half of the cases, we downgraded for important imprecision. Finally, for the group criteria related to “disease mongering”, we downgraded for overall high risk of bias.

Results by criteria group

We summarise the main findings here. Details of these analyses are reported in S10 File in the extended data.¹⁰⁴

Criteria related to associations and randomization

Many news reports included causal claims based on associations, and few said whether trials were randomized. We included three estimates, from three studies,^{19,32,45} related to the IHC Key Concepts that “*An outcome may be associated with a treatment but not caused by it*” (1.2d) and “*Comparison groups should be as similar as possible*” (2.1a). Based on estimates for single criteria with a low risk of bias, 51% of news reports make causal claims based on non-randomized studies (95% CI 40%-62%)³² and 9% specify whether a given trial was randomized (95% CI 7%-12%).¹⁹

Criteria related to the need for comparisons

Some news reports concluded that an intervention has an effect despite a lack of a comparator, and many did not include evidence to support the effectiveness of the intervention. We included three estimates, from three studies,^{12,32,38} related to the IHC key concept that “*Identifying effects of treatments depends on making comparisons*” (1.2f). Based on an estimate from a single study, three quarters (74%) of news reports avoid concluding there is a beneficial effect when there is a lack of a comparator (95% CI 63%-83%).³² Two studies found that only a third (34%)¹² to half (55%)³⁸ of news reports include evidence about the effects of the intervention (95% CIs 25%-43% and 38%-72%). The risk of bias for these two criteria was high.

Criteria related to consistency between studies

Few news reports considered and soundly assessed the consistency of evidence. We included a single criterion, from the study by Krauth and Apollonio,²⁸ related to the IHC Key Concept that “*The results of one study considered in isolation can be misleading*” (1.2g) and “*Consider how certain you can be about each advantage and disadvantage*” (3.3d). The risk of bias was low for this criterion. The study found that 13% (95% CI 7%-26%) of news reports published in American newspapers and magazines about tobacco cessation therapy “consider” and make a “well-founded” assessment of consistency between studies.

Criteria related to “disease mongering”

Some news reports committed disease mongering. We included five estimates, from five studies,^{12,13,26,38,40} related to disease mongering, which is addressed by the IHC Key Concept that “*Earlier detection of ‘disease’ is not necessarily better*” (1.2k). All five criteria had a high risk of bias. Overall, 86% of news reports do not commit disease mongering (95% CI 78%-91%) (Figure 3). The certainty of the evidence for that estimate is low due to substantial inconsistency ($I^2 = 92%$) and the high risk of bias.

Criteria related to conflicts of interest

Many news reports failed to address conflicts of interest. We included seven estimates, from seven studies,^{19,26,27,37,38,40,46} related to the IHC Key Concept that “*Competing interests may result in misleading claims*” (1.3b). All criteria had a low risk of bias. Overall, 22% (95% CI 7%-49%) in some way disclose potential conflicts of interest (Figure 3). The certainty of the evidence for that estimate is low, due to substantial inconsistency ($I^2 = 99%$) and important imprecision.

Criteria related to opinions

Some news reports failed to adequately distinguish opinion and fact. We included a single estimate, from the study by Krauth and Apollonio,²⁸ related to the IHC Key Concept that “*Opinions alone are not a reliable basis for claims*” (1.3d). The study evaluated American newspaper and magazine reports about tobacco cessation therapy. It found that three quarters (73%) of news reports adequately distinguish opinion and fact (95% CI 6%-83%). The criterion had a low risk of bias.

Criteria related to study design and risk of bias in general

Many news reports did not address the “quality of evidence”. We included five estimates, from five studies,^{13,26,28,38,40} related to the risk of bias (“type”, “credibility” or “quality” of the evidence), which is addressed by multiple low-level IHC Key Concepts under the high-level concepts “*Comparisons of treatments should be fair*” (2.1) and “*Syntheses of studies need to be reliable*” (2.2). Three of the criteria had a high risk of bias. The proportion of the news reports that meet the criterion used in each study varied from 8% (95% CI 3%-20%)²⁸ to 77% (95% CI 73%-81%).⁴⁰

Criteria related to blinding

Few news reports said whether trials were “double blinded”. We included a single criterion, from the study by Bubela *et al.*,¹⁹ related to the IHC Key Concepts that “*The people being compared should be cared for similarly apart from the treatments being studied*” (2.1c), “*If possible, people should not know which of the treatments being compared they are receiving*” (2.1d) and “*Outcomes should be assessed in the same way in all the groups being compared*” (2.1e). The study evaluated reports in newspapers in English-speaking countries about herbal remedy and conventional pharmaceutical trials. It found that 7% (95% CI 6%-9%) of news reports specify if a trial was “double-blinded”. The criterion had a low risk of bias. The term “double blind” is a problematic term since it can have several meanings.³⁷ The study did not address this.

Criteria related to loss to follow-up

Few news reports said anything about follow-up of participants in trials. We included a single criterion, from the study by Bubela *et al.*,¹⁹ related to the IHC Key Concept that “*It is important to assess outcomes in all (or nearly all) the people in a study*” (2.1g). This was the same study that evaluated whether news reports specify whether a trial was “double blind”.¹⁸ The risk of bias was low for this criterion. The study found that 7% (95% CI 5%-9%) of news reports specify withdrawals or dropouts from a trial.

Criteria related to selective reporting

Few news reports selectively reported “statistically significant” results. We included two estimates, from two studies,^{32,45} related to IHC Key Concepts that “*Failure to consider unpublished results of fair comparisons may result in estimates of effects that are misleading*” (2.2b) and “*Deeming results to be ‘statistically significant’ or ‘nonsignificant’ can be misleading*” (2.3g). Overall, 93% (95% CI 87%-97%) of news reports avoid selectively reporting “statistically significant” results (Figure 3). The certainty of the evidence for that estimate is high.

Criteria related to qualitative descriptions of effects

Many news reports used misleading language to describe intervention effects and did not quantify effects. We included seven estimates, from five studies,^{26,27,32,37,40} related to the IHC Key Concept that “*Verbal descriptions of the size of effects alone can be misleading*” (2.3a). Overall, four studies that reported six criteria related to quantification of effects found that 53% (95% CI 36%-69%) of news reports quantify effects, as opposed to only describing them qualitatively. There was a high risk of bias for two of the criteria. The certainty of the evidence for this estimate is also low, due to substantial inconsistency ($I^2 = 99%$) and important imprecision. Overall, two studies, found that 68% (95% CI 35%-90%) of news reports do not use misleading language to describe intervention effects. One of the criteria had a risk of bias. The certainty of the evidence for that estimate is low, due to substantial inconsistency ($I^2 = 98%$) and important imprecision.

Criteria related to absolute effects

Most news reports did not report absolute effects. We included five estimates, from five studies,^{12,13,37,38,47} related to the IHC Key Concept that “*Relative effects of treatments alone can be misleading*” (2.3b). All five criteria had a low risk of bias. Overall, 17% (95% CI 4%-49%) of news reports report absolute effects (Figure 3). The certainty of the evidence for that estimate is low, due to substantial inconsistency ($I^2 = 97%$) and important imprecision.

Criteria related to the play of chance

Few news reports claimed that an intervention has a benefit despite a small sample size, and more than half specified the sample size of trials, but few clearly and soundly “assessed the precision”⁸ of reported estimates of effect. We included four estimates, from four studies,^{19,28,32,36} related to the IHC Key Concepts that “*Small studies may be misleading*” (2.3d) and “*The use of p-values may be misleading; confidence intervals are more informative*” (2.3f). Based on estimates from single studies 95% (95% CI 88%-98%) of news reports do not claim there was a benefit despite a small sample size,³² and 63% (95% CI 59%-67%) specify the sample size of trials.¹⁹ Two studies found that 6% (95% CI 1%-38%)³⁶ and 8% (95% CI 3%-20%)²⁸ of news reports “clearly and soundly” assess the precision of estimates or risk of random error.

Criteria related to subgroup analyses

Most news reports did not focus on inappropriate subgroups. We included a single estimate, from the study by Yavchitz *et al.*,⁴⁵ related to the IHC Key Concept that “*Results for a selected group of people within a study can be misleading*” (2.3e). The study found that 94% (95% CI 85%-98%) of news reports avoid focusing on an “inappropriate subgroup”. The criterion had a low risk of bias.

Criteria related to lack of evidence

Most news reports avoided claiming that interventions have equivalent effects based on a difference that was not “statistically significant”. We included a single estimate, from the study by Yavchitz *et al.*,⁴⁵ related to the IHC Key Concept that “*Lack of evidence of a difference is not the same as evidence of ‘no difference’*” (2.3h). The study found that 93% (95% CI 83%-97%) of news reports do not claim equivalence when there was a “statistically nonsignificant” difference.⁴⁵ The criterion had a low risk of bias.

Criteria related to options

Most news reports did not include information about alternative interventions. We included seven estimates, from six studies,^{13,26,38,40,48} related to the IHC Key Concept “*Be clear about what the problem or goal is and what the options are*” (3.1a). Overall, 36% (95% CI 26%-47%) of news reports include information about alternative interventions (Figure 3). The certainty of the evidence for that estimate is moderate, due to substantial inconsistency ($I^2 = 96%$).

Criteria related to extrapolations

Some news reports focused on surrogate outcomes or extrapolated from surrogate to important outcomes. Some news reports did not report indications for interventions and most did not report contraindications. Few news reports did not clearly report to whom the information applied. Some extrapolated from study participants to a larger or different population. Many news reports did not consider the availability of the intervention or the dosage of the intervention. Some interventions extrapolated from the study intervention to a different intervention.

Altogether, we included 21 criteria, from 12 studies,^{12,13,19,26,28,32,35,38,40,45,48,49} related to the IHC Key Concepts that “*Attention should focus on all important effects of treatments, and not surrogate outcomes*” (3.2a), “*Fair comparisons of treatments in animals or highly selected groups of people may not be relevant*” (3.2b), “*The treatments compared should be similar to those of interest*” (3.2c) and “*There should not be important differences between the circumstances in which the treatments were compared and those of interest*” (3.2d). There was a low risk of bias for all the criteria.

We included two criteria, from two studies^{32,48} related to the IHC Key Concept that “*Attention should focus on all important effects of treatments, and not surrogate outcomes*” (3.2a). Overall, 82% (95% CI 76%-86%) of news reports do not solely focus on surrogate outcomes or extrapolate from surrogate to important outcomes. The certainty of the evidence for this estimate is high.

We included eight criteria, from six studies,^{28,32,35,38,48,49} related to IHC Key Concept that “*Fair comparisons of treatments in animals or highly selected groups of people may not be relevant*” (3.2b). Five of these criteria, from three of the studies,^{38,48,49} addressed reporting of indications and contraindications. Indications are reported in 70% (95% CI 64%-76%) to 97% (95% CI, 95%-98%) of news reports, whereas contraindications are only reported in 4% (95% CI 2%-8%) to 13% (95% CI 11%-16%) of news reports. Two criteria, from two studies,^{28,35} addressed whether it was clear to whom the information applies. These criteria are met in 94% (95% CI 81%-98%) to 100% (95% CI 98%-100%) of news reports. The other criterion, from the study by Haneef *et al.*,³² assessed whether results were extrapolated from study participants to a larger or different population. The study found that 85% (95% CI 72%-92%) of news reports do not extrapolate to a larger or different population.

We included eight criteria, from eight studies,^{13,19,26,32,38,40,49} related to the IHC Key Concept that “*The treatments compared should be similar to those of interest*” (3.2c). Five of these criteria addressed availability of the intervention. The availability of the intervention is reported in 41% (95% CI 37%-45%) to 79% (95% CI 66%-88%) of news reports. Two of the criteria addressed dosage. Those studies found that 14% (95% CI 11%-17%)¹⁹ to 50% (95% CI 46%-54%)⁴⁹ of news reports specify or mention dosage. The other criterion addressed extrapolation from the study intervention to a different intervention. That study³² found that 85% (95% CI 72%-92%) of news reports do not extrapolate to a different intervention.

Criteria related to animal studies

Some news reports extrapolated a beneficial effect from an animal study to humans. We included a single estimate, from the study by Haneef *et al.*,³² related to the IHC Key Concept that “*Fair comparisons of treatments in animals or highly selected groups of people may not be relevant*” (3.2b). The study, which had a sample of 29 news reports, found that 79% (95% CI 61%-90%) of reports do not extrapolate beneficial effects in humans from animal studies.³² The criterion had a low risk of bias.

Criteria related to pros and cons

Few news reports failed to mention at least one benefit of an intervention. On the other hand, few mentioned or discussed the cost of the intervention, and most did not mention or adequately discuss or explain potential harms.

Altogether, we included 32 estimates, from 17 studies,^{12,13,26–28,32,35–38,40,45–50} related to IHC Key Concepts that “*Treatments can cause harms as well as benefits*” (1.1a) and “*Weigh the benefits and savings against the harms and costs of acting or not*” (3.3a). One criterion was also included with criteria related to qualitative descriptions of effects, related to the IHC Key Concept that “*Verbal descriptions of the size of effects alone can be misleading*” (2.3a). That criterion was the only one that had a high risk of bias.

We included 14 criteria, from 13 studies,^{12,13,27,32,37,38,40,45–50} related to harms. Overall, 40% (95% CI 23%-61%) of news reports about the effects of health interventions mention or adequately discuss or explain potential harms of the intervention. The certainty of the evidence for that estimate is low, due to substantial inconsistency ($I^2 = 99\%$) and important imprecision.

We included 10 criteria, from nine studies,^{13,26,27,37,38,40,47–49} related to cost. Overall, 18% (95% CI 12%-28%) mention or discuss the cost of the intervention. The certainty of the evidence for that estimate is moderate, due to substantial inconsistency ($I^2 = 98\%$).

We included two criteria, from two studies,^{38,48} related to benefits. Overall, 97% (95% CI 56%-100%) of news reports mention at least one benefit. The certainty of the evidence for that estimate is moderate, due to imprecision.

Discussion

Key findings

Overall, we found that researchers have conducted many empirical studies of the quality of news reports about the effects of health interventions, assessing diverse samples of reports in terms of medium, country and categories of interventions, using many different tools and criteria. Some criteria did not make sense and some relevant aspects of quality were not covered. We were able to synthesis results for some criteria groups, showing that there were several prevalent and important problems with the quality of news reports about the effects of health interventions. Many of the reports gave an unbalanced and oversimplified picture of the potential consequences of the interventions, leaving out information about conflicts of interest, alternative interventions, potential harms, and costs, and failing to quantify effects in absolute terms or at all.

Specifically, we identified 44 primary studies for inclusion in the qualitative summary (see extended data - S1 Table¹⁰⁴). Newspaper was the most common medium. Of the 39 studies that specified a medium, newspaper content was explicitly included in 35 (90%). The United States (US) was the most common country in which content was published. Of the 40 studies where at least one country was specified, 12 (30%) included content published in the US, while 25 (63%) included content published in English-speaking countries. “Modern” medicine was the most common category of health interventions. In the 32 samples where a category of intervention was specified, 24 (75%) included news reports about “modern” medicine. No study specified whether sampled reports were published by commercial or non-commercial outlets. Most studies used a tool that was original or modified.

The Index of Scientific Quality was the most common tool (set of criteria). It was used in seven studies (16%).^{28,30,31,33–36} However, a larger number of studies used tools directly or indirectly based on the criteria developed by Moynihan *et al.*,³⁷ including the four studies that used the Media Doctor Australia tool^{12,13,38,39} and the three studies that used the Health News Review tool.^{25,26,40}

Of the 44 studies included in the qualitative summary, we included 18 (41%) in the meta-analyses. We mapped all 208 quality criteria used in those studies against the IHC Key Concepts. Some of the criteria were in conflict with IHC Key Concepts, such as those that asked whether research discussed in a news report was peer-reviewed. We were unable to identify or include quality criteria related to several highly relevant IHC Key Concepts, for example “*Average differences between treatments can be misleading*” (2.3c).

In the end, we included 108 of the 208 criteria (52%) in the meta-analyses, organised in 19 groups, as well as eight subgroups. We calculated overall estimates and rated the certainty of the evidence for five of the 19 main criteria groups (26%) and 6 of the 8 subgroups (75%). In those cases, we found:

- 86% of news reports commit “disease mongering” (95% CI 78%-91%) (low certainty),
- 22% adequately address conflicts of interest (95% CI 7%-49%) (low certainty),
- 93% do not selectively report “statistically significant” results (95% CI 87%-97%) (high certainty),
- 53% quantify effects (95% CI 36%-69%) (low certainty),
- 68% do not use misleading language (95% CI 35%-90%) (low certainty),
- 17% report absolute effects (95% CI 4%-49%) (low certainty),
- 36% include adequate information about alternative interventions (95% CI 26%-47%) (moderate certainty),
- 82% do not focus on surrogate outcomes or extrapolate from surrogate to important outcomes (95% CI 76%-86%) (high certainty),
- 40% mention, discuss or explain potential harms of the intervention (95% CI 23%-61%) (low certainty),
- 18% mention or discuss the cost of the intervention (95% CI 12%-28%) (moderate certainty), and
- 97% mention at least one benefit (95% CI 56%-100%) (moderate certainty).

Context

Media doctor Australia and health news review

Some of the studies included in our review report partial results from initiatives aimed at helping journalists and the public assess health information in the news and improving health journalism. This includes studies related to the Media Doctor Australia initiative^{12,13,39} and its American counterpart Health News Review (or HealthNewsReview.org).^{25,26} These initiatives involved an editorial team rating news reports about the effects of health interventions on a running basis and posting individual ratings with lay explanations on a dedicated website. Both have been shut down. We were unable to include the most complete published results from either initiative in our meta-analyses, but those results were consistent with the earlier results that we were able to include, as well as our overall findings.

In terms of Media Doctor Australia, we were able to include the studies by Smith *et al.*¹² and Bonevski *et al.*¹³ in the meta-analyses. The former includes results from the first 104 ratings, published between Feb. 1 and Sept. 1, 2004, while the latter includes results from the first 222 ratings of news reports about “complementary” and “alternative” medicine, published between Jan. 1, 2004, and Sept. 1, 2007. Meanwhile, Wilson *et al.* include results from 1230 ratings, posted between March 2004 and June 2008, in a paper titled “Media reporting of Health Interventions: Signs of Improvement, but Major Problems Persist”.³⁹ We included the study by Wilson *et al.* in our qualitative summary, but not in the meta-analyses, since the authors only reported mean scores and not the sample sizes for each quality criterion. We knew from the two other Media Doctor Australia studies that sample sizes varied between criteria because a criterion was not always considered relevant to a given news report and, in those cases, not applied. In other words, we could not extract the exact proportions of news reports satisfying each criterion from the study by Wilson *et al.* That said, as the title of paper suggests, there were no notable differences between the new Media Doctor Australia results and the previous results reported by Smith *et al.* (see [Table 2](#) in the study by Wilson and colleagues³⁹).

In terms of Health News Review, we were able to include the studies by Schwitzer²⁵ and Walsh-Childers *et al.*²⁶ in the qualitative summary. The former included results from 500 Health News Review ratings, published between April 2006 and February 2008, while the latter included results from 1889 ratings published from July 2005 through March 2013. Hence, we only included the study by Walsh-Childers *et al.* in the meta-analyses. The Health News Review team continued to publish ratings until the end of 2019, finishing at 2610.⁵¹ In a blog post published Dec. 20, 2018, the founder and publisher, Schwitzer, shared a final “report card” with some of the results from all 2610 ratings.⁵¹ Like with the Media Doctor Australia studies, there were no notable differences between the most recent results, published in Schwitzer's blog post, and those previously reported and included in this review, from the study by Walsh-Childers *et al.*

Underlying problems

Problems with the quality of information in health news reports could be explained by underlying problems occurring at different stages of the information pipeline, from when scientists plan, conduct and report studies, to when academic institutions and journals promote studies, to when news outlets cover studies. Not to mention deliberate misinformation by commercial industries and organisations.⁵²

Ioannidis has estimated that most published research findings are false due to small studies and small effect sizes, testing many relationships without preselecting those tests, flexibility in study designs, definitions, outcomes and analyses, financial and other conflicts of interest, and competition.⁵³ Similarly, Chalmers and Glasziou have estimated that 85% of research is wasted due to wrong research questions, unnecessary or poorly designed studies, failure to publish relevant research promptly or at all, and biased or unusable reports of research.⁵⁴ Meanwhile, Chi and colleagues have found problems with the quality of information in biomedical literature that correspond with problems found in health news reporting, for example inappropriate use of causal language.⁵⁵ A vast amount of poorly planned, conducted and reported research provides a terrible starting point for health and science journalism. Parenthetically, these studies provide additional evidence that peer review is a poor indicator of quality.

Several studies have identified problems with press releases about health research. Wang *et al.* found randomised trials were less likely to be the subject of a medical journal press release compared to observational studies.⁵⁶ Woloshin *et al.* assessed press releases from academic medical centres and found many were directly misleading or misleading by omission.⁵⁷ Yavchitz *et al.* found press releases for two-arm, parallel-group randomised trials often included some form of “spin”, for example conflating lack of evidence of a difference with no difference.⁴⁵ Sumner *et al.* have found exaggerations in press releases from both academic institutions⁵⁸ and scientific journals,⁵⁹ such as causal claims based on observational studies, were associated with exaggerations in related news reports. This finding was supported by a replication study conducted by Bratton *et al.*⁶⁰

Amend and Secko conducted a qualitative synthesis of studies exploring journalists' experience of covering health and science.⁶¹ They included 21 studies involving 788 journalists in total. Many of the themes cutting across studies were barriers to high-quality health and science journalism, including deadline pressure and limited time for researching and preparing stories; limited time or space for telling stories; limited budgets and staff, including a lack of specialised health and science journalists; competition and commercialization; pressure from advertisers and interest groups; and lack of education or training.

Other mass media sources of health information

Other studies have focused on other mass media sources of lay information about the effects of health interventions. This includes both other forms of journalism, besides news reports, and other mass media, besides news media. Some of those studies were included in this review. Overall, it appears quality varies depending on the source, but that there are problems across the board, granted we have not systematically reviewed or critically appraised studies focusing on other sources.

Newspaper features and advice columns – Heaner assessed news features and news reports about nutrition and physical activity.²⁹ She found features were generally of higher quality, which is unsurprising given that they were on average twice as long and probably had longer deadlines. Molnar *et al.* assessed medical advice columns for elderly readers and found that recommendations were often inappropriate or potentially dangerous.⁶²

Entertainment – Wilson *et al.* assessed the quality of advice popular magazine articles and found especially those with “health” in the title of the publication presented poor-quality, unreliable advice.⁶³ Korownyk *et al.* and Mishori *et al.* assessed recommendations made on medical talk shows.^{64,65} Both studies found the recommendations were often unreliable or misleading.

Social media – Moorhead *et al.* conducted a systematic review of the uses, benefits and limitations of social media for health communication and found that concern about the quality of information was common across studies. Li *et al.* found that over a quarter (19) of the 69 most-viewed YouTube videos about Covid-19 contained misleading information, including “inappropriate recommendations”, consistent with previous studies of YouTube videos about other pandemic diseases.⁶⁶ Haber *et al.* found many of the health science articles most often shared on Twitter and Facebook in 2015, as well as media coverage of those articles, often used causal language that was too strong for the strength of the evidence.²⁴

Advertising – There is a long tradition of misleading advertisements for health interventions. Dushman has provided 10 examples from the early 20th century, including an ad for Dr. William O. Coffee's “world-famous”, “marvelous” and “remarkable” treatment for deafness.⁶⁷ Frosch *et al.* and Faerber and Kreling have found problems with modern advertisements for both prescription and over-the-counter medicines.^{68,69} In sales materials for herbal dietary supplements, the United States Government Accountability Office has found “improper” and illegal claims that the supplements can cure cancer, among other conditions.⁷⁰

Other consumer and patient sources – In a systematic review of studies assessing the quality of online health information for consumers, Eysenbach *et al.* found 55 of 79 studies (70%) had concluded that quality was a problem.⁷¹ In more recently published studies, others have found problems with the information on anti-vaccination websites,⁷² cancer and oncological industry websites,⁷³ websites for clinics offering weight loss surgery,⁷⁴ and websites for fertility centres.⁷⁵ In their study of media coverage of practice-changing clinical trials in oncology, Andrew *et al.* found the cancer and industry websites provided higher-quality information than newspaper and cable news reports.⁷³

Studies focusing on patient sources of information have found problems with patient health portals,⁷⁶ patient materials,⁷⁷ and brochures from a health fair.³¹ Glenton and Oxman (AO) surveyed and interviewed leaders of organisations representing health care users.⁷⁸ They concluded that the organisations did not appear to promote evidence-based health care and that when they did promote scientific information, they appeared to do so uncritically, relying on limited sources and traditional authorities.

Last but not least, in a review published in 2019, Oxman and Paulsen assessed free, online sources of “trustworthy” information about the effects of health interventions.⁷⁹ They found that patients and the public could access trustworthy information using two websites: Cochrane Evidence (www.cochrane.org/evidence)⁸⁰ and Informed Health (www.informedhealth.org).⁸¹ However, they concluded that both websites could be improved by consistently reporting information about the size of both benefits and harms, and the certainty of the evidence. They noted that many websites excluded from their review claim to provide evidence-based or reliable information about the effects of health interventions, but that it was difficult to assess the reliability of the information on those websites since the information was not explicitly based on systematic reviews.

Other problems with health news

Other studies have found other problems with health news, besides the quality of information. Wang *et al.* found that observational studies and small randomised trials reporting surrogate outcomes generated as much news coverage as large, randomised trials reporting important outcomes.⁵⁶ Dumas-Mallet *et al.* found that newspaper reports about biomedical research were typically about findings from initial studies.⁸² Such findings were often contradicted by a meta-analysis, but this was rarely reported by news outlets.

Strengths

As far as we are aware, this is the first systematic review of the quality of any type of health news. It provides a larger, more detailed picture of the quality of health news than any primary study that we identified. It also provides a starting point for similar reviews and alternative analyses, and it highlights important issues for interpreting and planning new studies. We conducted an expansive search and identified several studies outside the peer-reviewed literature (see extended data - S1 Table¹⁰⁴). We assessed records and full texts in any language for inclusion in the review, and included studies in five different languages in the qualitative summary (English, German, Norwegian, Slovenian and Spanish) (see extended data - S1 Table¹⁰⁴). We synthesised results from up to 13 different studies, using different tools and criteria, and pooled results for up to 4116 unique news reports (see extended data - S8 and S10 Files¹⁰⁴). The review provides an accessible summary of what is known about the quality of news reports about the effects of health interventions, in print, broadcast and online news media, for those aspects of quality that have been measured and reported. It also identifies gaps in what is known and directions for future research.

Limitations

The findings of this review are limited to information about the effects of health interventions and “traditional” news media. In the section “Other mass media source of health information”, we reference and discuss findings from research focusing on other news media and other non-news media, but we have not systematically reviewed or critically appraised such studies. Furthermore, we did not explore applicability of the synthesised results to different types of traditional news media due to limited sample sizes, as well as inadequate reporting in the included studies.

Examples of other types of health information include the accuracy of diagnostic tests and information about the health effects of exposures. News reports about the health effects of exposures (e.g. fatty foods) can be difficult to distinguish from news reports about the effects of health interventions (e.g. a low-fat diet). Some primary studies included in this review conflated the two, such as the study by Hackman and Moe, which assessed news about nutrition-related research.⁸³ We included this study in the qualitative summary, but not the meta-analyses. We did this because key concepts for assessing or communicating information about the effects of interventions may not apply to the same degree, or at all, to information about associations between exposures and outcomes. For example, Woodruff and Sutton point out that randomised trials are “virtually precluded” from evidence about hazardous environmental exposures because of ethical considerations.⁸⁴ It is possible, if not likely, that there are important differences in quality between news reports about the effects of exposures versus interventions.

We excluded studies or criteria that were only descriptive, even though criteria that were descriptive sometimes provided the same data as criteria explicitly intended to assess quality. For example, Mercurio and Elliott describe Australian news about “alternative” medicine for cancer, but do not explicitly assess the quality of those news reports,⁸⁵ so we excluded the study. However, one of their coding variables was “Mention of risks, benefits and costs”. The data for this variable would have fit with data in our meta-analyses, in the group Criteria related to pros and cons. We consider it unlikely that inclusion of studies or criteria that were only descriptive would have substantially altered the findings of this review. Similarly, we only included studies in the meta-analyses if they were in English or Norwegian (one of the languages spoken fluently by the primary investigator, MO), but consider it unlikely that the studies reported in other languages would have substantially altered the findings of this review either, given that the underlying problems discussed earlier on are unrelated to language.

We did not code raw data ourselves or contact authors for help where data were inadequately reported for our purposes. Finally, we did not consider which criteria were more informative, only whether they were eligible and sensible, nor did we consider statistical measurement properties.⁸⁶

Implications

Implications for different stakeholders are summarised in [Table 5](#) and explained in more detail in the subsequent sections.

Table 5. Summary of implications.

Stakeholder		Implications
All		• Pay attention to relevant IHC Key Concepts
Journalism		
• Journalists		• Approach research and press releases critically • Search Cochrane Evidence and Informed Evidence for relevant summaries
• Editors • News outlets		• Train health and science journalists • Provide adequate conditions for researching, preparing and telling stories
Research		
• Researchers • Journals • Academic institutions • Funders	General	• Reduce false research findings and research waste
• Investigators • Peer reviewers	Quality of health news	• Think carefully about the goal, tools and quality criteria • Consider applicability, explanatory factors and risk of bias • Be explicit, clear and comprehensive in reporting
Consumers and patients		• Approach health news critically • Search Cochrane Evidence and Informed Evidence for relevant summaries
Citizens and policymakers		• Advocate for teaching IHC Key Concepts in school

Implications for journalists, editors, and news outlets

There is clearly room for improving the quality of health and science journalism. Based on our meta-analyses, journalists and editors should pay special attention to the following IHC Key Concepts (see extended data - S6 File¹⁰⁴) when communicating information about the effects of health interventions:

- “Treatments can cause harms as well as benefits” (1.1a),
- “Competing interests may result in misleading claims” (1.3b),
- “Verbal descriptions of the size of effects alone can be misleading” (2.3a),
- “Relative effects of treatments alone can be misleading” (2.3b),
- “Be clear about what the problem or goal is and what the options are” (3.1a), and
- “Weigh the benefits and savings against the harms and costs of acting or not” (3.3a).

Based on individual estimates from primary studies included in this review (see extended data - S8 and S10 Files¹⁰⁴), there is likely also room for improvement related to other highly relevant IHC Key Concepts, such as “*The treatments compared should be similar to those of interest*” (3.2c). Moreover, there is likely room for improvement related to highly relevant IHC Key Concepts that were not captured by any quality criteria, such as “*Average differences between treatments can be misleading*” (2.3c) (see extended data - S10 File¹⁰⁴).

Journalists and editors should take a critical approach to reports of research, especially single studies and observational studies, and to press releases (see “Underlying problems”). When reporting on the effects of interventions, they can search the Cochrane Evidence (www.cochrane.org/evidence)⁸⁰ and Informed Health (www.informedhealth.org)⁸¹ websites for summaries of relevant systematic reviews. If possible, news outlets should provide or support special training for health and science reporters, as well as provide adequate time for researching and preparing stories, and adequate time or space to tell stories.

Journalists and editors who are interested in learning more about IHC Key Concepts can do so for free on the “That’s a claim!” website (www.thatsclaim.org),⁸⁷ where all of the concepts are explained in plain language, and there are references to relevant learning resources for each individual concept. They can also find free learning resources for respective concepts via the Teachers of Evidence-Based Health Care portal (www.teachingebhc.org).⁸⁸

Implications for researchers, scientific journals, academic institutions, and research funders

Researchers, journals, and academic institutions should pay special attention to the same IHC Key Concepts as journalists and editors, as well as other highly relevant concepts (see previous section), when reporting the results of studies assessing the effects of interventions in research papers and press releases, as well other contexts, such as conferences. More fundamentally, they should work to reduce false findings and research waste. This also goes for organisations that fund research.

The results of this review have a series of implications for investigators planning new studies of the quality of information about the effects of health interventions in mass media (as well as peer reviewers assessing new studies), from objectives, to methods, to reporting. First, they should think carefully about their goal. New studies can be designed to have practical value, for example informing interventions to improve a particular population’s ability to think critically about health information by identifying the IHC Key Concepts that are most relevant for that population—e.g. a study of health claims in advertisements targeted at young people. In addition, they should consider research gaps that we have revealed. Again, there does not appear to be studies assessing the quality of health news reports in terms of several highly relevant IHC Key Concepts (see extended data - S6 Table¹⁰⁴). Neither does there appear to be studies that compare the quality of reports published by commercial outlets versus public or independent outlets. Others have pointed out that research is needed about how best to communicate information about the effects of health interventions,⁴³ including how to communicate uncertainty.⁸⁹

Second, they should think carefully about what tools and quality criteria to use, whether using a previously developed tool, adapting one, developing a new tool, or some combination. They should avoid criteria that are overly general (e.g. global scores and composites) or unclear (e.g. criteria assessing whether “double blinding” is mentioned), and criteria that do not make sense (e.g. criteria assessing whether research was peer-reviewed). They should be deliberate about the trade-offs between how straightforward it is to apply a criterion and report the results versus how informative it is. For example, it is relatively easy to see whether a news report mentions the design of a study. It is more difficult, but also more informative, to assess whether the report adequately discusses strengths and limitations of said design.

Third, regarding samples of news reports, they should consider applicability and potential explanatory factors, for example those we highlighted in the protocol for this review (see extended data - S1 File¹⁰⁴) as possible variables for subgroup analyses: medium; time period; financial model (commercial vs. non-commercial); specialisation (specialised health or science journalist vs. other); country income level; and newspaper type (broadsheet vs. tabloid). Fourth, they should consider risk of bias. If possible, two researchers should assess each news report or other unit of mass media content, and the researchers should be blinded to factors that could introduce bias, such as the name of the journalist or publication. This is especially important for criteria that require substantial judgement.

Fifth, when reporting new studies, researchers should be explicit, clear and comprehensive. They should include the study objective, the sampling frame, and the selection criteria and technique, with concern for applicability and comparing results across studies, including results for subgroups. They should include the time period in which the content was published, as well as when it was assessed; the country in which it was published or, if online content, the country that the outlet primarily targets or where it is based; the categories of interventions that the content focuses on; and the financial models of the outlets. If using a previously developed tool, they should provide reference to where the development is described, note whether they made any adjustments and, if so, explain them. If using a new tool, they should clearly describe how it was developed and what informed it, possibly in a separate paper. If the researchers received any training in applying the criteria, they should describe what this entailed. Finally, they should report results for each quality criterion separately, not just mean scores, and report absolute numbers (the total number of news reports or other units to which the criterion was applied and the number of units that satisfied the criterion), including for subgroups.

Implications for consumers and patients

The implications for consumers and patients are similar to those for journalists and editors. They should pay special attention to the same IHC Key Concepts, as well as other highly relevant concepts (see “Implications for journalists”); they should approach health information in the mass media critically; they can refer to the Informed Choices and

Cochrane Evidence websites for summaries of relevant systematic reviews; and they can take advantage of the “That’s a claim!” website and Teachers of Evidence-Based Health Care portal to learn more about IHC Key Concepts.

Implications for citizens and policymakers

The results of this review, as well as the results of related studies referenced in the section “Other mass media sources of health information”, show that the global “infodemic” was compounded by Covid-19, but started long before the World Health Organization coined the term in their situation report on the virus published Feb. 13, 2020.³ The infodemic is so dangerous because we know so many people are unable to critically assess claims or evidence about the effects of health interventions.^{90,91} When people act on those claims or beliefs that are unreliable or fail to act on reliable advice, they may suffer unnecessarily or waste resources. It follows that infodemic is probably an important factor in the enormous, worldwide overuse of ineffective or harmful interventions⁹² and underuse of effective interventions.⁹³ Even one misleading news report might contribute to wasteful, harmful, and even deadly choices. A dramatic example of misleading health news is coverage of the baseless suggestion that the measles, mumps and rubella vaccine causes autism.⁹⁴ Incidentally, news reports often focus on the benefits of interventions while ignoring or downplaying the harms, but this example illustrates that unbalanced reporting can also go in the opposite direction.

If the only goal of news about the effects of health interventions was to facilitate well-informed health choices, the content and format might look more like Cochrane plain language summaries, which have been systematically developed towards exactly said end.^{95,96} The reality is that news outlets existentially depend on getting people’s attention, not informing them. Commercial media rely on advertising, not least advertisements for health interventions, which provides incentive to sensationalise. In turn, this can create pressure on competing public and independent outlets to do the same. Along the same lines, the *raison d’être* of press releases is inciting media coverage, while scientists are forced to “publish or perish”. Setting aside systemic barriers to reliable health news (see the section “Underlying problems”), it may be difficult if not impossible for journalists to produce news reports that satisfy all sensible, applicable criteria identified in this review in a report that people would understand and want to read, not to mention other potential criteria related to other highly relevant IHC Key Concepts. Even if a report includes and appropriately presents all relevant information about an intervention, it is possible that many people will not read more than the headline, missing important details, such as potential side effects.

It is possible for news outlets, including commercial ones, to improve their health and science coverage without shrinking their audiences and perhaps even increasing their numbers of readers, listeners, or viewers in some cases. But putting an end to unreliable health news is a Sisyphean task, not to mention health misinformation from other sources. In 2009, during a debate titled “Does the media support or sabotage health?”, Goldacre suggested: “The future of people getting information about health does not lie in journalism but people having direct access to information or the world of blogs where people can link direct to primary sources. There is no role for the health journalists of the future other than as entertainers”.⁹⁷ However, it seems unlikely that giving people more health information—even if it is reliable—will solve the problem, when they are already overwhelmed and lack the skills needed to “separate the wheat from the chaff”.

As the saying goes, give someone a fish and you will feed them for a day; teach them to fish and you feed them for a lifetime. In other words, you can provide people with reliable health information or advice, and maybe they will believe you. If you help them understand why it is reliable, or why a claim is unreliable, you are preparing them to assess information and make informed choices for themselves in the future, inoculating them against unreliable claims and uninformed choices. Hence, we should be helping people gain basic critical appraisal and decision-making skills, in addition to providing access to reliable information. Moreover, by improving people’s ability to assess evidence or recognise a lack of reliable evidence, the public may start to demand reliable studies, reducing false findings and research waste.

Focusing on the education system allows for reaching large groups of people, gathered specifically to learn. It is important to start young, when people have time to reinforce and build upon what they have learned, and they have less to unlearn. The Informed Health Choices network, which includes authors of this review, has developed respective interventions to help primary school children and university students master select IHC Key Concepts, and has started to develop an intervention for secondary school students as of writing (www.informedhealthchoices.org/secondary-school-resources).⁹⁸

Educational interventions, like health intervention should be rigorously evaluated to find their effects, both positive and negative.^{99–101} The IHC team evaluated the primary school intervention in a cluster-randomised trial in Uganda, involving more than 10,000 children. The results showed that even children in a low-income setting, with large class

sizes (on average 84 children in both intervention and control schools at the beginning of the trial), can learn to apply IHC Key Concepts, and the children retained what they learned for at least a year.¹⁰² Importantly, the intervention was developed using a human-centred design approach.¹⁰³ A process evaluation linked to the trial showed that children, teachers, headteachers and parents were positive to the project and supported expanding it to other populations.¹⁰⁴

The Norwegian *Bak overskriftene* (“Behind the headlines”) intervention targeted at university students takes advantage of health claims in the mass media.¹⁰⁵ In the first instance, it has been implemented in a mandatory, introductory course on evidence-based health care, for students in the first two years of all health sciences bachelor programmes at Oslo Metropolitan University (about 1600 students per year). The intervention exploits that media content by design is simple, familiar, relatable, and entertaining, making it a favourable starting point for learning about critical appraisal compared to scientific literature, which typically includes jargon and excludes narrative. Journalism school students and staff have helped develop the intervention, which could potentially be used to train both journalism students and professionals, thereby increasing the quality of health and science news, assuming the intervention is properly evaluated and found effective.

Finally, mastering IHC Key Concepts might help people assess information and make informed choices about non-health interventions as well. In collaboration with colleagues from other disciplines, the IHC team has established that most of the concepts apply to at least 13 other fields: agriculture, economics, education, environmental management, international development, informal learning, management, nutrition, planetary health, policing, speech and language therapy, social welfare, and veterinary medicine.¹⁰⁰

Data availability

Underlying data

All data underlying the results are available as part of the article and its supporting information files and tables.

Extended data

Zenodo: Quality of news reports about the effects of health interventions - Supporting Information. <http://doi.org/10.5281/zenodo.4781182>.¹⁰⁴

This project contains the following extended data:

- **S1 File. Protocol.** “Quality of news media reports about the effects and costs of health interventions: Systematic review protocol”.
- **S2 File. Protocol Appendix 1.** “Appendix 1: Reference list of potentially eligible studies”.
- **S3 File. Protocol Appendix 2.** “Appendix 2: Description of search for an existing review”.
- **S4 File. Dissertation.** “Criteria used to measure the quality of news media reports about the effects of health interventions: Systematic review”.
- **S5 File. Detailed search strategy.**
- **S6 File. Informed Health Choices Key Concepts.** 2019 version.
- **S7 File. Detailed information about deviations from the protocol.**
- **S8 File. Forest plots for individual criteria groups.**
- **S9 File. Individual risk of bias assessments.**
- **S10 File. Individual criteria included in meta-analyses.** Sample characteristics, verbatim and reworded criteria, related IHC Key Concepts, overall risk of bias, estimates and confidence intervals, and sample sizes.
- **S1 Table. Included studies.** References for all studies included in qualitative summary and reason for exclusion from meta-analyses or number of criteria included in meta-analyses.

- **S2 Table. Sample characteristics and tools by study.** Sample size, medium/media, country/countries, time period(s), intervention category/categories, and tool(s).
- **S3 Table. Objectives of included studies.**
- **S4 Table. Sampling frames and methods.**
- **S5 Table. Reported subgroup analyses.**
- **S6 Table. Number of quality criteria included in the meta-analysis per relevant IHC Key Concept.**

Reporting guidelines

Zenodo: PRISMA checklist for 'Quality of information in news media reports about the effects of health interventions: systematic review and meta-analyses'. <http://doi.org/10.5281/zenodo.4781182>.¹⁰⁴

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

Author contributions

MO drafted the protocol and this paper. AO provided feedback on the draft protocol. All co-authors provided feedback on drafts of this paper. MO, LL and AQ piloted the eligibility criteria for primary studies. MO conducted all database searches, including citation searches. MO, LL, GG, DA, AQ and KB checked reference lists. All authors apart from CR contributed to study selection. All authors apart from AO and CR contributed to data extraction. MO and AO assessed, mapped and grouped the quality criteria. CR conducted the meta-analyses. MO and LL assessed risk of bias. MO and AO rated the certainty of the evidence.

Acknowledgements

Mike Clarke provided early guidance in a course on systematic reviews. Carl Heneghan and Kamal Mahtani were MO's MSc supervisors and provided feedback on the protocol and the first part of the review (MO's dissertation). Marit Johansen provided feedback on the search strategy. Anette Blüme, Tina Poklepović Perićić, Rob J.P.M. Scholten and Michel Wensing helped assess full texts in German, Slovenian and Dutch for inclusion in the qualitative summary. We are grateful to Mark Oette for providing a copy of his study, included in the qualitative summary. This review is dedicated to Lisa Schwartz who was a pioneer in the field¹⁰⁶ and a dear colleague and friend.

References

1. Zuger A: **A Journal Stands Out in Prestige and Longevity.** *The New York Times*. Mar 2012; **20**: 4.
[Reference Source](#)
2. Smith FA, Trivax G, Zuehlke DA, et al.: **Health Information during a Week of Television.** *N Engl J Med*. 1972; **286**: 516–520.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. World Health Organization: **Novel Coronavirus(2019-nCoV): Situation Report - 13.** 2020 Feb.
[Reference Source](#)
4. Chew C, Eysenbach G: **Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak.** *PLoS One*. 2010; **5**: e14118.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. Oxman M: **Quality of news media reports about the effects and costs of health interventions: Systematic review protocol.** 2018.
[Reference Source](#)
6. Moher D, Liberati A, Tetzlaff J, et al.: **Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement.** *PLoS Med*. 2009; **6**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Moberg J, Austvoll-Dahlgren A, Treweek S, et al.: **The plain language Glossary of Evaluation Terms for Informed Treatment choices (GET-IT) at www.getitglossary.org.** *Res All*. 2018; **2**: 106–121.
[Publisher Full Text](#)
8. Oxman AD, Guyatt GH, Cook DJ, et al.: **An index of scientific quality for health reports in the lay press.** *J Clin Epidemiol*. 1993; **46**: 987–1001.
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Zeraatkar D, Obeda M, Ginsberg JS, et al.: **The development and validation of an instrument to measure the quality of health research reports in the lay media.** *BMC Public Health*. 2017; **17**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Bosch F, Escalas C, Forteza A, et al.: **A checklist for improving drug information in the general press: the importance of reporting on the phases and uncertainty of research.** *Revista Española Comun en Salud*. 2018; **9**: 203–214.
[Publisher Full Text](#)
11. Ouzzani M, Hammady H, Fedorowicz Z, et al.: **Rayyan—a web and mobile app for systematic reviews.** *Syst Rev*. 2016; **5**: 210.
[Publisher Full Text](#)
12. Smith DE, Wilson AJ, Henry DA: **Monitoring the quality of medical news reporting: early experience with media doctor.** *Med J Aust*. 2005; **183**: 190–193.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Reference Source](#)
13. Bonevski B, Wilson A, Henry DA: **An Analysis of News Media Coverage of Complementary and Alternative Medicine.** *PLoS One*. 2008; **3**: e2406.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Bolarinwa OA: **Principles and methods of validity and reliability testing of questionnaires used in social and health science researches.** *Niger Postgrad Med J*. 2015; **22**: 195.
[PubMed Abstract](#) | [Publisher Full Text](#)

15. Oxman A, Chalmers I, Dahlgren A, *et al.*: **Key Concepts for Informed Health Choices: A framework for enabling people to think critically about health claims.** 2019; [Reference Source](#)
16. Chalmers I, Oxman AD, Austvoll-Dahlgren A, *et al.*: **Key Concepts for Informed Health Choices: a framework for helping people learn how to assess treatment claims and make informed choices.** *BMJ Evidence-Based Med.* 2018; **23**: 29–33. [PubMed Abstract](#) | [Publisher Full Text](#)
17. Oxman AD, García LM: **Comparison of the Informed Health Choices Key Concepts Framework to other frameworks relevant to teaching and learning how to think critically about health claims and choices: a systematic review [version 1; peer review: 3 approved].** *F1000Res.* 2020; **9**. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Austvoll-Dahlgren A, Oxman AD, Chalmers I, *et al.*: **Key concepts that people need to understand to assess claims about treatment effects.** *J Evid Based Med.* 2015; **8**: 112–125. [PubMed Abstract](#) | [Publisher Full Text](#)
19. Bubela T, Boon H, Caulfield T: **Herbal remedy clinical trials in the media: A comparison with the coverage of conventional pharmaceuticals.** *BMC Med.* 2008; **6**. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Guyatt GH, Oxman AD, Vist GE, *et al.*: **GRADE: what is “quality of evidence” and why is it important to clinicians?** *BMJ.* 2008; **336**: 995–998. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Cochrane Effective Practice and Organisation of Care (EPOC): **Reporting the effects of an intervention in EPOC reviews.** *EPOC Resources for review authors.* 2017. [Reference Source](#)
22. Wilson AJ: **Assessing the Quality of Health News Stories in the Australian Media Using the Media Doctor Website.** *PhD thesis.* 2010. [Reference Source](#)
23. Haber N, Smith ER, Andrews K, *et al.*: **Causal strength and language in health academia and in media sources (CLAIMS): a systematic review of academic health literature in social media.** [Reference Source](#)
24. Haber N, Smith ER, Moscoe E, *et al.*: **Causal language and strength of inference in academic and media articles shared in social media (CLAIMS): A systematic review.** *PLoS One.* 2018; **13**: e0196346. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. Schwitzer G: **How do US journalists cover treatments, tests, products, and procedures? An evaluation of 500 stories.** *PLoS Med.* 2008; **5**: 0700–0704. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Walsh-Childers K, Braddock J, Rabaza C, *et al.*: **One Step Forward, One Step Back: Changes in News Coverage of Medical Interventions.** *Health Commun.* 2018; **33**: 174–187. [PubMed Abstract](#) | [Publisher Full Text](#)
27. Høye S, Hjortdahl P: **«Ny vidunderpille!» - Hva skriver norske aviser om nye legemidler?** *Tidsskr den Nor Lægeforening.* 2002; **122**: 1671–1676.
28. Krauth D, Apollonio D: **Accuracy of popular media reporting on tobacco cessation therapy in substance abuse and mental health populations.** *BMJ Open.* 2015; **5**: 1–7. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Heaner MK: **Media Messages about Nutrition and Physical activity: A content analysis of the accuracy and reporting characteristics of newspaper coverage in the U.S.** 2009.
30. Halvorson RT, Stewart CC, Thakur A, *et al.*: **Scientific Quality of Health-Related Articles in Specialty Cannabis and General Newspapers in San Francisco.** *J Health Commun.* 2018; **23**: 993–998. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Johansen LW, Bjørndal A, Flottorp S, *et al.*: **An evaluation of health information in newspapers and brochures: What should one believe?** *Tidsskr Nor Lægeforen.* 1996; **116**: 260–264. [PubMed Abstract](#)
32. Haneef R, Lazarus C, Ravaud P, *et al.*: **Interpretation of Results of Studies Evaluating an Intervention Highlighted in Google Health News: A Cross-Sectional Study of News.** *PLoS One.* 2015; **10**: e0140889. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. Biondo E, Khoury MC: **Validation of a questionnaire to assess the quality of health information in Argentinian newspapers.** *Biomédica.* 2005; **25**: 366–376. [PubMed Abstract](#) | [Reference Source](#)
34. Ferraz AV, Llop Rius R, Bosch Ferrer M, *et al.*: **Noticias sobre medicamentos en los suplementos de salud de la prensa Española.** *Med Clin (Barc).* 2005; **124**: 754–755. [Publisher Full Text](#)
35. Huey J, Apollonio DE: **A content analysis of popular media reporting regarding increases in minimum ages of legal access for tobacco.** *BMC Public Health.* 2018; **18**: 0–7. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Lewis M, Orrock P, Myers S: **Uncritical reverence in CM reporting: Assessing the scientific quality of Australian news media reports.** *Heal Sociol Rev.* 2010; **19**: 57–72. [Publisher Full Text](#)
37. Moynihan R, Bero L, Ross-Degnan D, *et al.*: **Coverage by the news media of the benefits and risks of medications.** *N Engl J Med.* 2000; **342**: 1645–1650. [PubMed Abstract](#) | [Publisher Full Text](#)
38. Robinson C: **Media Reporting of Health Interventions in New Zealand: A Retrospective Analysis.** 2018. [PubMed Abstract](#) | [Publisher Full Text](#)
39. Wilson A, Bonevski B, Jones A, *et al.*: **Media Reporting of Health Interventions: Signs of Improvement, but Major Problems Persist.** *PLoS One.* 2009; **4**: e4831. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Stassen W: **Health research as news in South Africa: measuring the quality of health journalism at six daily newspapers.** 2016.
41. Neumeyer-Gromen A, Bodemer N, Müller SM, *et al.*: **Ermöglichen Medienberichte und Broschüren informierte Entscheidungen zur Gebärmutterhalskrebsprävention? Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz.** 2011; **54**: 1197–1210. [Publisher Full Text](#)
42. Jefferson T, Alderson P, Wager E, *et al.*: **Effects of editorial peer review: A systematic review.** *JAMA.* 2002; **287**: 2784–2786. [PubMed Abstract](#) | [Publisher Full Text](#)
43. Oxman AD, Glenton C, Flottorp S, *et al.*: **Development of a checklist for people communicating evidence-based information about the effects of healthcare interventions: a mixed methods study.** *BMJ Open.* 2020; **10**: e036348. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
44. Moynihan R, Heath I, Henry D: **Selling sickness: the pharmaceutical industry and disease mongering.** *BMJ.* 2002; **324**: 886–891. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
45. Yavchitz A, Boutron I, Bafeta A, *et al.*: **Misrepresentation of Randomized Controlled Trials in Press Releases and News Coverage: A Cohort Study.** *PLoS Med.* 2012; **9**: e1001308. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
46. Colombo C, Mosconi P, Buratti MG, *et al.*: **Press coverage of hormone replacement therapy and menopause.** *Eur J Obstet Gynecol Reprod Biol.* 2010; **153**: 56–61. [PubMed Abstract](#) | [Publisher Full Text](#)
47. Iaboli L, Caselli L, Filice A, *et al.*: **The Unbearable Lightness of Health Science Reporting: A Week Examining Italian Print Media.** *PLoS One.* 2010; **5**: e9829. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
48. Cassels A, Hughes MA, Cole C, *et al.*: **Drugs in the news: An analysis of Canadian newspaper coverage of new prescription drugs.** *CMAJ.* 2003; **168**: 1133–1137. [PubMed Abstract](#) | [Free Full Text](#)
49. Finer D, Tomson G: **Medicines in the news: Content analysis of drug articles in four Vietnamese newspapers.** *Int J Risk Saf Med.* 1999; **12**: 109–122. [Reference Source](#)
50. Woloshin S, Schwartz LM: **Media reporting on research presented at scientific meetings: More caution needed.** *Med J Aust.* 2006; **184**: 576–580. [PubMed Abstract](#) | [Publisher Full Text](#)
51. Schwitzer G: **A final HealthNewsReview.org report card from 3,200+ systematic reviews of health care news stories & PR releases.** 20 Dec 2018 [cited 5 Feb 2021]. [Reference Source](#)
52. Goldberg RF, Vandenberg LN: **The science of spin: targeted strategies to manufacture doubt with detrimental effects on environmental and public health.** *Environ Heal.* 2021; **20**: 33. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
53. Ioannidis JPA: **Why Most Published Research Findings Are False.** *PLoS Med.* 2005; **2**. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
54. Chalmers I, Glasziou P: **Avoidable waste in the production and reporting of research evidence.** *Lancet.* 2009; **374**: 86–89. [Publisher Full Text](#)
55. Chiu K, Grundy Q, Bero L: **‘Spin’ in published biomedical literature: A methodological systematic review.** *PLoS Biol.* 2017;

- 15: e2002173.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
56. Wang MTM, Bolland MJ, Gamble G, *et al.*: **Media coverage, journal press releases and editorials associated with randomized and observational studies in high-impact medical journals: A cohort study.** *PLoS One.* 2015; **10**: 1–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
57. Woloshin S, Schwartz LM, Casella SL, *et al.*: **Press releases by academic medical centers: Not so academic?** *Ann Intern Med.* 2009; **150**: 613–618.
[PubMed Abstract](#) | [Publisher Full Text](#)
58. Sumner P, Vivian-Griffiths S, Boivin J, *et al.*: **The association between exaggeration in health related science news and academic press releases: retrospective observational study.** *BMJ.* 2014; **349**: g7015–g7015.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
59. Sumner P, Vivian-Griffiths S, Boivin J, *et al.*: **Exaggerations and Caveats in Press Releases and Health-Related Science News.** *PLoS One.* 2016; **11**: e0168217.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
60. Bratton L, Adams RC, Challenger A, *et al.*: **The association between exaggeration in health-related science news and academic press releases: a replication study.** *Wellcome Open Res.* 2019; **148**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
61. Amend E, Secko DM: **In the Face of Critique: A Metasynthesis of the Experiences of Journalists Covering Health and Science.** *Sci Commun.* 2012; **34**: 241–282.
[Publisher Full Text](#)
62. Molnar FJ, Man-Son-Hing M, Dalziel WB, *et al.*: **Assessing the quality of newspaper medical advice columns for elderly readers.** *CMAJ.* 1999; **161**: 393–395.
[PubMed Abstract](#) | [Free Full Text](#)
63. Wilson A, Smith D, Peel R, *et al.*: **A quantitative analysis of the quality and content of the health advice in popular Australian magazines.** *Aust N Z J Public Health.* 2017; **41**: 256–258.
[PubMed Abstract](#) | [Publisher Full Text](#)
64. Korownyk C, Kolber MR, McCormack J, *et al.*: **Televised medical talk shows - What they recommend and the evidence to support their recommendations: A prospective observational study.** *BMJ.* 2014; **349**: 1–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
65. Mishori R, Weinfeld J, Lau L, *et al.*: **Verifying and Assessing Content of Recommendations Made on Daytime Medical Talk Shows.**
66. Li HO-Y, Bailey A, Huynh D, *et al.*: **YouTube as a source of information on COVID-19: a pandemic of misinformation?** *BMJ Glob Heal.* 2020; **5**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
67. Dushman A: **Ads and labels from early 20th-century health fraud promotions.** *AMA J Ethics.* 2018; **20**: 1082–1093.
[PubMed Abstract](#) | [Publisher Full Text](#)
68. Frosch DL, Krueger PM, Hornik RC, *et al.*: **Creating Demand for Prescription Drugs: A Content Analysis of Television Direct-to-Consumer Advertising.** *Ann Fam Med.* 2007; **5**: 6–13.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
69. Faerber AE, Kreling DH: **Content Analysis of False and Misleading Claims in Television Advertising for Prescription and Nonprescription Drugs.** *J Gen Intern Med.* 2013; **29**: 110–118.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
70. United States Government Accountability Office: **Herbal Dietary Supplements: Examples of Deceptive or Questionable Marketing Practices and Potentially Dangerous Advice.** 2010.
[Reference Source](#)
71. Eysenbach G, Powell J, Kuss O, *et al.*: **Empirical Studies Assessing the Quality of Health Information for Consumers A Systematic Review.** *JAMA.* 2002; **287**: 2691–2700.
[PubMed Abstract](#) | [Publisher Full Text](#)
72. Wolfe Robert M, Sharp Lisa K, Lipsky MS: **Content and design attributes of antivaccination web sites.** *JAMA.* 2002; **287**: 3245–3248.
[PubMed Abstract](#) | [Publisher Full Text](#)
73. Andrew P, Vickers MM, O'Connor S, *et al.*: **Media Reporting of Practice-Changing Clinical Trials in Oncology: A North American Perspective.** *Oncologist.* 2016; **21**: 269–278.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
74. Groven KS, Braithwaite J: **Happily-ever-after: Personal narratives in weight-loss surgery advertising.** *Health Care Women Int.* 2016; **37**: 1221–1238.
[PubMed Abstract](#) | [Publisher Full Text](#)
75. Spencer EA, Mahtani KR, Goldacre B, *et al.*: **Claims for fertility interventions: A systematic assessment of statements on UK fertility centre websites.** *BMJ Open.* 2016; **6**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
76. Glenton C, Paulsen EJ, Oxman AD: **Portals to Wonderland: health portals lead to confusing information about the effects of health care.** *BMC Med Inform Decis Mak.* 2005; **5**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
77. Coulter A, Entwistle V, Gilbert D: **Informing patients: an assessment of the quality of patient information materials.** London: King's Fund; 1998.
78. Glenton C, Oxman AD: **The use of evidence by health care user organizations.** *Heal Expect.* 1998; **1**: 14–22.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
79. Oxman AD, Paulsen EJ: **Who can you trust? A review of free online sources of "trustworthy" information about treatment effects for patients and the public.** *BMC Med Inform Decis Mak.* 2019; **19**: 35.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
80. Cochrane: **Cochrane Evidence.** [cited 18 Feb 2021].
[Reference Source](#)
81. The Institute for Quality and Efficiency in Health Care: **Informed Health.**
82. Dumas-Mallet E, Smith A, Boraud T, *et al.*: **Poor replication validity of biomedical association studies reported by newspapers.** *PLoS One.* 2017; **12**: e0172650.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
83. Hackman EM, Moe GL: **Evaluation of newspaper reports of nutrition-related research.** *J Am Diet Assoc.* 1999; **1564**–1566.
[PubMed Abstract](#) | [Publisher Full Text](#)
84. Woodruff TJ, Sutton P: **An Evidence-Based Medicine Methodology To Bridge The Gap Between Clinical And Environmental Health Sciences.** *Health Aff.* 2011; **30**: 931–937.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
85. Mercurio R, Elliott JA: **Trick or treat? Australian newspaper portrayal of complementary and alternative medicine for the treatment of cancer.** *Support Care Cancer.* 2011; **19**: 67–80.
[PubMed Abstract](#) | [Publisher Full Text](#)
86. Mookink LB, Prinsen CA, Patrick DL, *et al.*: **COSMIN manual for systematic reviews of PROMs, user manual.** 2018: 1–78.
[Reference Source](#)
87. Centre for Informed Health Choices, InfoDesignLab, Epistemonikos Foundation: **That's a claim!** [cited 18 Feb 2021].
[Reference Source](#)
88. International Society for Evidence-Based Health Care: **Teachers of Evidence-Based Health Care.** [cited 18 Feb 2021].
[Reference Source](#)
89. Büchter RB, Betsch C, Ehrlich M, *et al.*: **Communicating Uncertainty in Written Consumer Health Information to the Public: Parallel-Group, Web-Based Randomized Controlled Trial.** *J Med Internet Res.* 2020; **22**: e15899.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
90. Boutron I, Haneef R, Yavchitz A, *et al.*: **Three randomized controlled trials evaluating the impact of "spin" in health news stories reporting studies of pharmacologic treatments on patients'/caregivers' interpretation of treatment benefit.** *BMC Med.* 2019; **17**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
91. Dahlgren A, Furuseth-Olsen K, Rose CJ, *et al.*: **The Norwegian public's ability to assess treatment claims: results of a cross-sectional study of critical health literacy.** *F1000Res.* 2020; **9**: 179.
[Publisher Full Text](#)
92. Brownlee S, Chalkidou K, Doust J, *et al.*: **Evidence for overuse of medical services around the world.** *Lancet.* 2017; **390**: 156–168.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
93. Glasziou P, Straus S, Brownlee S, *et al.*: **Evidence for underuse of effective medical services around the world.** *Lancet.* 2017; **390**: 169–177.
[PubMed Abstract](#) | [Publisher Full Text](#)
94. Goldacre B: *Bad science.* London: Fourth Estate; 2008.
95. Santesso N, Rader T, Nilsen ES, *et al.*: **A summary to communicate evidence from systematic reviews to the public improved understanding and accessibility of information: a randomized controlled trial.** *J Clin Epidemiol.* 2015; **68**: 182–190.
[PubMed Abstract](#) | [Publisher Full Text](#)
96. Cochrane Community: **Cochrane begins project to improve Plain Language Summaries.** 18 May 2020 [cited 18 Feb 2021].
[Reference Source](#)
97. Coombes R: **Health journalism: two clicks away from Britney?** *BMJ.* 2009; **338**.
[Publisher Full Text](#)
98. Rosenbaum S, Oxman M, Oxman A, *et al.*: **Human-centred design development of Informed Health Choices (IHC) learning resources for secondary school students: Protocol.** 2019.
[Reference Source](#)

99. Sharples JM, Oxman AD, Mahtani KR, *et al.*: **Critical thinking in healthcare and education.** *BMJ.* 2017; **357**: j2234.
[PubMed Abstract](#) | [Publisher Full Text](#)
100. Aronson JK, Barends E, Boruch R, *et al.*: **Key concepts for making informed choices.** *Nature.* 2019; **572**: 303–306.
[PubMed Abstract](#) | [Publisher Full Text](#)
101. Zhao Y: **What works may hurt: Side effects in education.** *J Educ Chang.* 2017; **18**: 1–19.
[Publisher Full Text](#)
102. Nsangi A, Semakula D, Oxman AD, *et al.*: **Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects, 1-year follow-up: A cluster-randomised trial.** *Trials.* 2020; **21**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
103. Nsangi A, Semakula D, Rosenbaum SE, *et al.*: **Development of the informed health choices resources in four countries to teach primary school children to assess claims about treatment effects: A qualitative study employing a user-centred approach.** *Pilot Feasibility Stud.* 2020; **6**: 18.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
104. Nsangi A, Semakula D, Glenton C, *et al.*: **Informed health choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: Process evaluation.** *BMJ Open.* 2019; **9**: e030787.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
105. Oxman M, Habib L, Jamtvedt G, *et al.*: **Using claims in the media to teach essential concepts for evidence-based healthcare.** *BMJ Evidence-Based Med.* 2020; **0**: 1–3.
[PubMed Abstract](#) | [Publisher Full Text](#)
106. Roberts S: **Dr. Lisa Schwartz, Critic of Medical Excess, Is Dead at 55.** In: *The New York Times [Internet].* 6 Dec 2018 [cited 16 Feb 2021].
[Reference Source](#)

Open Peer Review

Current Peer Review Status:    

Version 1

Reviewer Report 23 August 2021

<https://doi.org/10.5256/f1000research.56222.r90284>

© 2021 Henry D. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **David Henry** 

Institute for Evidence-Based Healthcare, Bond University, Gold Coast, Australia

The other reviewers of this article have highlighted its considerable strengths and I won't repeat their comments. This is a solid attempt to summarise studies that have made structured attempts to quantify and report key characteristics of medical news stories that report therapeutic claims. It is very useful to have this summary.

Many of these studies were carried out to highlight the weaknesses in the reporting of medical advances in the mainstream media in the hope that this would raise standards. But there is no strong evidence that this has resulted from these efforts.

The studies included in this review preceded the news coverage of COVID-19, its preventive and therapeutic interventions. It would be very instructive to measure how the media performed during the epidemic and whether reporting evolved. My informal impression is that respect for science, and possibly the reports themselves, have improved. However, this has been countered by ignorant and some willfully and politically motivated inaccurate reporting in a few mainstream outlets and social media. My point is that the media landscape may have changed because of Covid-19.

I only have a few specific comments:

1. I don't think that conventional meta-analysis produces meaningful data in this situation. Meta-analysis is at its best when we can assume that there is an underlying single effect of interest that is measured imprecisely by individual studies. For instance, the relative effects of many important medications (e.g, statins, anticoagulants), which seem constant across different settings and background risks. Meta-estimates of prevalence are less helpful as there is usually an assumed variation. That is the case here where the study methods, story topics, settings, and reporters vary. There is no reason to assume a common effect to be meta-analysed. Some of the I^2 values are huge. In my view it would be preferable to present more as a scoping review, displaying and discussing the range of values reported by the studies.

2. The investigators, appropriately, have mapped the varying scales and items to a common framework. They have used their own Informed Health Choices (IHC) Key Concepts framework. That's fine, but the paper seems to assume a familiarity with the framework that many readers will not possess. I think a simplified version of the framework needs to be presented in the main text. At present, it seems to be buried in Supplementary data.
3. In the Abstract, the authors conclude: 'There is room for improving health news, but it is logically more important to improve the public's ability to critically appraise health information and make judgements for themselves.' The paper does not provide data to support this conclusion. The authors have done important ground-breaking work with IHC but the extensive coverage they provide themselves in the Discussion of an already long paper seems excessive. A separate editorial/opinion piece making these points would be welcome.
4. I think there may be a mistake in the dot point summary in the Discussion which states '• 86% of news reports commit "disease mongering" (95% CI 78%-91%) (low certainty).' In the paper they state 'Overall, 86% of news reports do not commit disease mongering (95% CI 78%-91%)'. I think the latter is correct.

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

Partly

Are the conclusions drawn adequately supported by the results presented in the review?

Partly

Competing Interests: I am a co-author of several studies included in this review.

Reviewer Expertise: clinical epidemiology, pharmacoepidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 09 August 2021

<https://doi.org/10.5256/f1000research.56222.r90288>

© 2021 Wilson A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Amanda Wilson 

School of Nursing and Midwifery, University of Technology Sydney, Sydney, NSW, Australia

This is a systematic review of studies assessing the quality of health news articles which provides a meta-analysis of the results. This is an interesting paper which provides a quantitative approach to what is likely to be considered a difficult area to quantify – that is quality of reporting. The influence of news reporting on health behaviours is well known and the current climate of pandemic and vaccination information makes this a highly pertinent area to explore.

The methodology is sound and rigorous with the authors following the gold standard approach of registering the review and publishing the protocol. All steps of the literature review are described in detail and the analyses and interpretation are sound.

The paper is extremely well written and I have little critical feedback on the structure of the research or presentation of the paper. My only criticism is that it is extremely long and possibly overly detailed. The length and detail are worthy of a Cochrane review but without the saving grace of a plain language summary. I wonder if some ruthless editing might enhance the number of readers who make it the whole way through. It feels like a PhD chapter rather than a journal article. Having said that, in the Discussion on page 18, I was interested to know which criteria did not make sense and which aspects of quality were not covered.

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

Yes

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Public health, health literacy, health reporting in the media, health education

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 04 August 2021

<https://doi.org/10.5256/f1000research.56222.r90763>

© 2021 Lexchin J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Joel R. Lexchin 

¹ School of Health Policy & Management, Faculty of Health, York University, Toronto, Canada

² Department of Family and Community Medicine, Faculty of Medicine, University of Toronto, Toronto, Canada

³ Department of Emergency Medicine, University Health Network, Toronto, Canada

This is an extremely comprehensive and well-done systematic review on an important topic. The authors have obviously put an enormous amount of effort into the project and the results show that. The study is well written and easy to follow. However, there are some areas where I believe that improvements can be made.

1. There is a potentially relevant article that the authors appear to have overlooked: Cassels *et al.* Open Medicine 2008;2(1):e20–23. Another study (Cassels *et al.*, 2003) is a longer version of one that the authors use – reference 48.
2. Page 3: The authors use a 2009 paper to claim that most people get health information either directly or indirectly from traditional news media as opposed to material that appears only on social media. However, studies that evaluate information sourced from the two different types of media may no longer be relevant to the present-day environment.
3. Page 18: The authors say that their findings come from a diverse range of countries but don't they mean primarily western countries?
4. Page 20: The length of a newspaper story is not necessarily a measure of quality. Cassels *et al.* have shown that there are few significant differences in the overall quality between short news briefs and longer stories. (Cassels *et al.*, 2003)
5. Page 22: The recommendations for improving the reporting of journalists, editors and news outlets at least partially assume the existence of resources and time to educate journalists and editors about science journalism. However, the authors do not assess whether the necessary resources are actually present and available especially in smaller media outlets.
6. Page 22: Informed Health is only available in English and German. Therefore, it is limited in terms of usefulness to those speaking other languages. Cochrane Evidence is superior in that regard.
7. Page 23 (last line): Is it Informed Choices or Informed Health? Are these the same or different?
8. Page 24: Earlier the authors refer to Informed Health and Informed Choices. Now there is a third name Informed Health Choices. Are all three the same?
9. Page 24: Have any of the educational interventions that the authors mention (references 99-104) looked at the long-term (e.g., more than 1 year) effects?

References

1. Cassels A, Lexchin J: How well do Canadian media outlets convey medical treatment information?: Initial findings from a year and a half of media monitoring by Media Doctor Canada. *Open Med.* 2008; **2** (2): e45-8 [PubMed Abstract](#)
2. Cassels A, Hughes M, Cole C, Mintzes B, et al.: Drugs in the News:How well do Canadian newspapers report the good,the bad and the ugly of new prescription drugs?. *Canadian Centre for Policy Alternatives.* 2003. [Reference Source](#)

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

Yes

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

Competing Interests: In 2017-2020, Joel Lexchin received payments for being on a panel at the American Diabetes Association, for talks at the Toronto Reference Library, for writing a brief in an action for side effects of a drug for Michael F. Smith, Lawyer and a second brief on the role of promotion in generating prescriptions for Goodmans LLP and from the Canadian Institutes of Health Research for presenting at a workshop on conflict-of-interest in clinical practice guidelines. He is currently a member of research groups that are receiving money from the Canadian Institutes of Health Research and the Australian National Health and Medical Research Council. He is a member of the Foundation Board of Health Action International and the Board of Canadian Doctors for Medicare. He receives royalties from University of Toronto Press and James Lorimer & Co. Ltd. for books he has written. He confirms that this did not interfere with his ability to do a fair and unbiased review of the manuscript.

Reviewer Expertise: Pharmaceutical policy, intellectual property rights

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 29 July 2021

<https://doi.org/10.5256/f1000research.56222.r89219>

© 2021 Moynihan R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Raymond Moynihan

Institute for Evidence-Based Healthcare, Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia

General Comments.

This review is extremely valuable – the first summarizing evidence about the quality of news reports about health interventions.

I have a few short general comments and some specific comments.

In relation to the use of the word “qualitative” - I wonder if the authors might consider using a different word – or adding more explanation – perhaps saying “qualitative summary of results” – as it sounds to the casual reader like when you refer to “qualitative” you may have done a SR of “qualitative” studies of news coverage (i.e. focus groups with journalists etc.) – whereas I think you have a different meaning in mind. (e.g. P8 “qualitative summary”, which is also mentioned in Abstract).

Re: IHC. Given the importance of IHC concepts for this SR, and the role of the authors in IHC world, and the advocacy of IHC in the Discussion, I think the IHC concepts etc. needs to be mentioned in the Introduction and expanded on in the Methods (as to how and why exactly the IHC concepts interacted with your analysis in the SR) - and the authors link to IHC can be mentioned in Methods somewhere. This is important and will help explain how the Discussion section becomes a little like advocacy for IHC materials.

The paper is long and complex – but with much valuable information. If any more material can be moved into Supp files, that will be helpful.

I have no biostats training so I have not assessed the stats.

Specific comments.

- Search strategy: It seems almost 2 years since it was updated. Authors might consider updating before indexing.
- P5: “We excluded inappropriate criteria...” Perhaps “ineligible” might be better?
- P7: Re: your sentences “Furthermore, we reframed negative quality criteria as positive, and paraphrased all quality criteria, for the data to be consistently framed and worded, and for the criteria to be succinctly labelled. For example, Bubela et al. report the number of news reports that “did not” specify whether a study was randomized, which was reframed as the number of news reports that did specify whether a study was randomized.”

While this makes sense, I presume that sometimes that re-framing was not straight forward – and I wonder if there is anything more to say on any complexities associated with this approach. Also, if you can give a little more detail in the Bubela example, with numbers, that could help the reader.

- P7: “Risk of Bias”. Can the authors add a few words in text explaining why the “risk of bias” was solely about the risk of detection bias? A reader of an SR might normally expect a few different domains, so it would be good to just explain why this single domain was most

relevant.

- P10: Table 2. I think Table 2 needs re-thinking – as it took me some time to work out what was going on. I was treating each row as a discreet coherent entity, but in fact you have multiple parts to each row. I am sure you can make it clearer somehow.
- P11: “In contrast, a checklist for people communicating evidence-based information about the effects of healthcare interventions includes the item “Help your audience to avoid misinterpreting continuous outcome measures,” granted the checklist is not for journalists.” I don’t fully understand meaning of this sentence - perhaps be a little clearer?
- P12: This is a structural suggestion. By the sub-heading “Synthesis of Results” I am really wanting to hear some of the key findings – in terms of content (i.e. a strong compelling prose version of the data in Figure 3) – but I feel I am still getting a lot of complex fine detail about the characteristics of studies and elements of them. I understand there is a lot to report – but some clear statements of the key findings – even summarised and abstracted a little – would be helpful around here to orient the reader.

Also – a related point - I find the three paragraphs on P12 (and into P13) complex and hard to follow. I wonder whether this level of detail is necessary in the main text – and whether it might be moved to a supplementary – but that is for authors to decide. At the very least – if there is any way to make it all a little clearer around this point – that would be great for the reader. An example of the mention of the “ninth criteria” from Bubela and 12th from Haneef. I think this is way too much detail for main text and authors run the risk of losing readers, for what is a very valuable SR.

Now I have read further I realise you have a long section - starting on P14 – on the results – as they relate to IHC concepts. After having read that very long section, I still feel the need for some kind of punchy summary of the key findings (e.g. harms, COIs, costs, other options not often mentioned...) somewhere in the results section. This kind of summary doesn’t appear until start of Discussion – but I think could be included earlier.

- P18: I wonder if the dot points at bottom of P18 might be better in the Results – and help to drive the kind of summary I am suggesting above.
- Discussion: the paper feels like it morphs in the Discussion into a passionate advocacy for IHC materials – with a sometimes moralistic tone. I think the tone could be addressed. And I also think that in order to justify the IHC material in the Discussion, more needs to be said about IHC in the Introduction – and the authors relations with IHC need to be made more explicit in the Methods – as noted above.

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are the conclusions drawn adequately supported by the results presented in the review?

Partly

Competing Interests: RM reports having collaborated closely with AO, off and on for roughly 10 years, more than 10 years ago. In the last 10 years RM reports having had minimal contact with AO, and minimal contact with MO, consisting of offering advice, such as reviewing documents, via brief emails, relating to projects undertaken by MO or AO. I understand this may be perceived as a potential conflict of interest, but I have endeavoured to conduct my brief review of this paper as fairly and objectively and in as unbiased a manner as I can.

Reviewer Expertise: Medicine and Media; Conflicts of Interest; Overdiagnosis

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research