PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The association of industry ties with outcomes of studies examining the effect of wholegrain foods on cardiovascular disease and mortality: Systematic review and Meta-analysis
AUTHORS	Chartres, Nicholas; Fabbri, Alice; McDonald, Sally; Turton, Jessica; Allman-Farinelli, Margaret; McKenzie, Joanne; Bero, Lisa

VERSION 1 - REVIEW

REVIEWER	Marion Nestle New York University, NY, USA I know the authors personally, have worked with them, and think they do excellent and important work
REVIEW RETURNED	29-Mar-2018

GENERAL COMMENTS	The authors (whom I know personally, have worked with, and respect and admire) pulled 22 studies out of 6818 for their analysis of industry-funding bias in studies of wholegrains and clinical outcomes. Studies funded by industry were slightly more likely to produce positive outcomes. Biases were common regardless of funding source. Nearly half the investigators failed to disclose conflict of interest. This study adds to the still quite small literature on funding bias in nutrition research. Specific comments 88. Say what the recommendations are. 91. Do you doubt the conclusions? Why suspect bias? A stronger rationale for choosing wholegrains would help (for example, studies of wholegrains and cancer risk produce null findings). file:///C:/Users/Marion%20Admin/Dropbox/DOCUMENTS/Foods/Grains/Makarem_WholeGrainsCA_NutrRev_16.pdf 92. A definition of wholegrains seems needed here. It would help to define wholegrain in terms of species (wheat, oats, etc), components (endosperm, bran, germ), and percentages (100%). The industry-sponsored Whole Grain Council says “A grain is considered to be a whole grain as long as all three original parts — the bran, germ, and endosperm — are still present in the same proportions as when the grain was growing in the fields” (https://wholegrainscouncil.org/whole-grains-101/whats-whole-grain-refined-grain). This includes ground up seeds, but what about flour (endosperm) to which bran and germ are later added? Or endosperm to which some proportion of whole grain is added? 304 ff. This section is confusing: “All studies contained a sponsorship disclosure statement,” but “ten studies did not contain a disclosure statement.” Clarification of this and subsequent statements in this paragraph is needed. 312. To my mind, wholegrain means 100%. Did nonindustry studies use that? Why 25%? Does this mean that industry studies called breakfast cereals wholegrain even though they may

	have been only 25% wholegrain? Or that 100% wholegrain cereals only comprised 25% of the total? This would be easier to understand if there were definitions given earlier. And this paragraph is hard to follow. It would help to have all sentences follow the same format starting with industry (or nonindustry) but not mixing them up. 332. Can you give an example of a confounder? 418. I am guessing that most people doing studies of the health benefits of wholegrains would have a vested interest (industry) or an ideological interest (nutritionists who believe wholegrains are especially nutritious and desirable). I would not expect industry funding to make much difference in this instance. 464. The one clear finding of this research is the nondisclosure. This needs even more emphasis since it's impossible to do this kind of research if disclosure doesn't happen.

REVIEWER	Giuseppe Grosso NNEdPro Global Center for Nutrition and Health, UK
REVIEW RETURNED	08-Apr-2018

GENERAL COMMENTS	The present study investigated the potential association between industry sponsorship and outcomes in studies on whole grains and CVD outcomes. The study is undoubtedly well conducted, with excellent methodology and proper execution. What I am concerned about are the conclusions. Based on results, the authors should "relax" a bit the concluding remarks, as - in fact - there was no such big evidence of bias in results, rather lack of statement for COI which, if not mandatory by journal policy, is often forgot by the authors. We know that the highest risk of bias affects RCT rather than cohort studies, and major drive to publish and ideally find significant results in cohort studies is for scientific reasons (and, being cynical, to improve personal background), but results from cohort studies are hard to affect (I don't see researcher inventing results or subverting common methods to try to obtain favorable results). In conclusions, authors should relax the conclusions of both abstract and text, authors may say "we detected potential risk of bias etc.", but quantitative results were not significant so makes no sense saying "they are in that way DESPITE not significant" (after all, heterogeneity between studies is very low, so the example of whole grains does not fit optimally to search for potential influence of industry on results). Notes on lack of COI are ok. A final short note, I don't have personal conflict of interest in providing this revision, it is solely my opinion.

REVIEWER	Eva María Navarrete-Muñoz CIBER of Epidemiology and Public Health, Spain
REVIEW RETURNED	01-Jun-2018

GENERAL COMMENTS	The study addresses an important issue about the possible influences of the sponsorship in the studies published about the relationship between whole grain and cardiovascular disease. This

	is a well written manuscript and focused on an interesting topic for BMJ Open's readers. However, there are some imprecision and some concerns about the methodology used that should be clarified. a) The most important concern is about the search strategy. It is should be more specific to respond of the aim of the study, for example, the search included a randomized control trial or intervention but the aim of the review should have only included observational studies. b) This review was conducted including observational studies (see lines 130-131), thus why did the authors decide to search in Cochrane Library? c) In the line 146, the authors said this review included healthy children but it is impossible they found children studies using outcome measurements defined in the lines 160-165. d) The abstract is so confused and it should be rewritten. In the results, they showed critical risk of bias but this concept was not previously defined. The aim is more clearly described in the article than in the abstract. e) In the background, we added information of previous studies which evaluate the role of sponsorship in studies with food. I recommended to read this publication Bes-Rastrollo et al. Financial conflicts of interest and reporting bias regarding the association between sugar-sweetened beverages and weight gain: a systematic review of systematic reviews. PLoS Med. 2013 Dec;10(12):e1001578; f) In lines 192-195, the authors said that effect size was defined as the risk ratio but in some cases they used hazard ratios to present the results. Moreover, in the case-control studies, the adequate association measure is the odds ratio not risk ratio. The authors should change this sentence. g) The same changes of the previous point should be conducted in the lines 272-273. h) In the data collection and analysis, the author mentioned an intervention period although the review only included observational studies. i) In line 283, the authors said that they undertook random effects meta-analysis to compare the association across subgroups. However, they should use random effects when I2>50% while fixed effects were observed in the rest of the cases. j) In the table 1, some categories should be reduced in order to avoid some collapse, especially in risk bias assessment. k) The results should be reorganized, firstly they should include the answer to primary aim and after the answer to secondary aim.

REVIEWER	Jonathan R Treadwell ECRI Institute, USA
REVIEW RETURNED	02-Nov-2018

GENERAL COMMENTS	The association of industry sponsorship with outcomes of studies examining the effect of wholegrain foods on cardiovascular disease and mortality: Systematic review and Meta-analysis To assist the BMJ, my comments are divided into Major and Minor Major comments Strict vs lenient usage of $p<0.05$. The authors suggest that in order for one of their included studies to legitimately conclude that there is an association between industry sponsorship and a positive effect of wholegrain foods, the study must have found $p<0.05$. Thus, they strictly used $p<0.05$ when examining the results of individual studies (with $p<0.05$ indicating favorable to industry, and $p>0.05$ indicating results unfavorable to industry). However, when describing their own results, they were less strict, and despite $p>0.05$ in several instances, allowed themselves a gray-area conclusion of "Studies with industry ties were more likely to have favourable results than those with no industry ties…although the result was not statistically significant". This choice in how to present their results ("more likely" despite $p>0.05$) makes the reader think they might have been biased to find the association for which they were looking. A more even-handed wording would have been something like "Data were inconclusive with respect to the association between industry ties and favorable results, as the relative risk could be as high as 2.35 or as low as 0.88". So the authors appear to be guilty of the same "discordance" between results and conclusions that they had planned to analyzed in one of their secondary outcomes (if any studies had had such a discordance, which they did not). The authors might instead consider a Bayesian analysis of the results, and simply state that given the data, there is an X% chance (with X being perhaps 80%) that studies with industry ties are more likely to have favorable results. That way, your words can be faithful to the data, and you're not a slave to an arbitrary 0.05 cutoff, and also you avoid any impression of bias towards a positive association. BMJ reviewers should understand that the first analysis mentioned in the abstract, the one with RR=1.44, is based simply on 2x2 table where 8 of 9 industry-ties studies had a favorable finding, as compared to 8 of 13 non-industry-ties studies had a favorable finding. So it's just 89% vs 62%. This is a very low power statistical test, since each study is simply put into one of four boxes in the 2x2 table…"vote-counting"…this low power perhaps explains the wide confidence interval. If 89% vs 62% had been based on more studies, such as 44 instead of 22, the same percentage difference would have been classified as statistically significant (16/18 vs 16/26 yields $p=0.045$). A better statistical approach, which the authors have also done, is to use each study's quantitative effect size, and see if the weighted average of industry-ties-studies results was any different from the weighted average of non-industry-ties-studies' results. This is called subgroup analysis. For this analysis, the authors also found statistical non-significance, but they chose to describe their results in a very different way as "we did not find an important difference…". This is puzzling, because in both cases, there is a nonsignificant effect size that is in the direction (at least) of

	suggesting that industry ties are more likely to find favorable results. So why the completely different interpretation? Using a Bayesian approach throughout (and associated Bayesian wording for conclusions such as there is an X% chance that a hypothesis is true) may resolve this rather glaring inconsistency. I would say that for the subgroup analysis, rather than presenting separately each subgroup's summary effect size, you should present the statistical contrast between subgroups (its point estimate and CI). This output is not in Figure 3 or 4, but the output does give the exact p value ("Test for subgroup difference chi2=0.46 p=0.50). Figure out the corresponding point estimates and their CI's, and present that. Regarding your wording in the abstract "These findings suggest, but do not establish…". This is very problematic wording. At heart, by the conventional standard for making scientific claims i.e. $p<0.05$, your results are inconclusive, and "inconclusive" should be your first word describing them. Notice that you felt that for the subgroup analyses you could firmly say "there is no important difference". Why couldn't the pro-wheat industry apply the same wording to your RR=1.44 finding due to the wide CI (ie they could call it no important difference)? You need to find a way to be consistent when interpreting different outcomes. I didn't see any clear delineation in your methods or results for, if confronted with $p>0.05$, you would conclude 1 "no important difference" vs 2 you would say it's inconclusive vs 3 you would use some complicated wording such as "suggests but do not establish". Minor comments The title should say "industry ties" instead of "industry sponsorship". This is because authors defined (and analyzed) industry ties as including EITHER industry sponsorship OR industry COI's amongst authors. In the abstract, within the results, it's confusing that you've flipped the meaning of being RR>1 vs RR<1. When you present RR=1.44, clearly you defined RR>1 as meaning that industry ties are associated with more favorable results, but later when you present RR=0.77 and RR=0.85, you are then defining RR>1 as meaning the opposite. Please make these consistent. I do think the latter set of analyses are much more important, as they take advantage of quantitative effect sizes, rather than just put each study into one of four boxes. Line 185 specify whether 0.05 had to be one tailed or two tailed. You might consider replacing "favorable" with "pro-industry" and "unfavorable" with "anti-industry". This might be a bit more direct and easier for readers. Line 253. "non-profit organizations" What if there's a non-profit organization that is anti-wheat? Or a pharma company that is anti-wheat? It's too simplistic to put all of these into your "industry" category. I had thought, when reading the intro/methods that all your "industry" were prescreened to be pro-wheat industry. So these needs to be more obvious to readers that you've combined industries in this way.

	Hazard ratios are A LOT like relative risks. You might consider doing a larger meta-analysis that includes both (and so perhaps ALL studies could be in it, for a new Figure 3). After all you included RR's and OR's in the same meta-analysis, and those are technically different (albeit similar given the low event rates).

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Marion Nestle

Institution and Country: New York University, NY, USA Please state any competing interests or state 'None declared': I know the authors personally, have worked with them, and think they do excellent and important work

The authors (whom I know personally, have worked with, and respect and admire) pulled 22 studies out

of 6818 for their analysis of industry-funding bias in studies of wholegrains and clinical outcomes.

Studies funded by industry were slightly more likely to produce positive outcomes. Biases were

common regardless of funding source. Nearly half the investigators failed to disclose conflict of interest.

This study adds to the still quite small literature on funding bias in nutrition research.

Specific comments

88. Say what the recommendations are.

Response: We have now included a sentence in line 91, p 5 'The guidelines conclude that there is a probable association between whole grain consumption and a reduced risk of cardiovascular disease'.

91. Do you doubt the conclusions? Why suspect bias? A stronger rationale for choosing wholegrains

would help (for example, studies of wholegrains and cancer risk produce null findings).

Response: We have added a sentence to the first para (Line 95, p5). "However, the beneficial effects of wholegrains on CVD when assessed in randomised controlled trials (RCTs) are uncertain".

92. A definition of wholegrains seems needed here. It would help to define wholegrain in terms of

species (wheat, oats, etc), components (endosperm, bran, germ), and percentages (100%). The industry sponsored Whole Grain Council says "A grain is considered to be a whole grain as long as all three

original parts — the bran, germ, and endosperm — are still present in the same proportions as when the

grain was growing in the fields" (https://wholegrainscouncil.org/whole-grains-101/whats-whole-grainrefined-grain). This includes ground up seeds, but what about flour (endosperm) to which bran and

germ are later added? Or endosperm to which some proportion of whole grain is added?

Response: Thank you for this comment (line 92 is a paragraph break). As the reviewer points out, there are multiple ways to define a wholegrain.

We have added text and references in line 99, p5 “Wholegrain products can be defined in various ways, including by the species (e.g., wheat, oats), components (e.g., endosperm, bran, germ), and percentages (e.g., 25%-100%). While some food regulators use a definition of 100% retention of wholegrain content, the epidemiological literature typically uses 25% or more retained content. In the development of the Australian Dietary Guidelines, the most common definition for wholegrain foods was those containing 25% or more of wholegrains.”

In our methods (line 162, p8) we note “We included studies that defined wholegrains in any way, as defined by the author.”

304 ff. This section is confusing: “All studies contained a sponsorship disclosure statement,” but “ten studies did not contain a disclosure statement.” Clarification of this and subsequent statements in this paragraph is needed.

Response: We report on funding disclosure statements and author conflict of interest statements separately. We have now changed lines 319, p16 to “Ten studies did not contain an author conflict of interest disclosure statement”.

312. To my mind, wholegrain means 100%. Did nonindustry studies use that? Why 25%? Does this mean that industry studies called breakfast cereals wholegrain even though they may have been only 25% wholegrain? Or that 100% wholegrain cereals only comprised 25% of the total? This would be easier to understand if there were definitions given earlier. And this paragraph is hard to follow. It would help to have all sentences follow the same format starting with industry (or non industry) but not mixing them up.

Response: As shown In Table 1, pp17, both industry (n=6) and non-industry sponsored studies (n=4) defined whole grains as >25%. As shown in supplementary file 4, these studies were either examining breakfast cereals on their own (n=2) or as part of total whole grain intake. As we were not answering a clinical question and not trying to understand the effect of various types of wholegrain, we state (line 162,p 8) “We included studies that defined whole grains in anyway, as defined by the author of the included study”.

To make this paragraph easier to follow, we have now amended line 326, p16 to follow the same format starting with industry “Industry sponsored studies were less likely (56%) to have a serious or critical risk of bias in classification of exposures than non-industry sponsored studies (85%)”.

332. Can you give an example of a confounder?

Response: We have added to line 342, p18 “For example, a confounder was fruit and vegetable intake. If this was not appropriately controlled for when assessing the effect of wholegrain intake on a cardiovascular disease outcome, the study was rated as having a risk of bias for confounding.”

418. I am guessing that most people doing studies of the health benefits of wholegrains would have a vested interest (industry) or an ideological interest (nutritionists who believe wholegrains are especially

nutritious and desirable). I would not expect industry funding to make much difference in this instance.

Response: Thank you for this comment.

464. The one clear finding of this research is the nondisclosure. This needs even more emphasis since it's impossible to do this kind of research if disclosure doesn't happen.

Response: We agree with this comment and an additional sentence has been added in the abstract and discussion:

Abstract (Line 54, p3) "Our findings support international reforms to improve the disclosure and management of conflicts of interest in nutrition research. Without such disclosures, it will not be possible to determine if the results of nutrition research are free of food industry influences and potential biases."

Manuscript (line 480, p24) "This research further strengthens calls for stricter policies relating to the disclosure and management of conflicts of interest in nutrition research. Without such disclosures, it will not be possible to determine if the results of nutrition research are free of food industry influences and potential biases."

Reviewer: 2

Reviewer Name: Giuseppe Grosso

Institution and Country: NNEdPro Global Center for Nutrition and Health, UK Please state any competing interests or state 'None declared': None declared

The present study investigated the potential association between industry sponsorship and outcomes in studies on whole grains and CVD outcomes. The study is undoubtedly well conducted, with excellent methodology and proper execution. What I am concerned about are the conclusions. Based on results, the authors should "relax" a bit the concluding remarks, as - in fact - there was no such big evidence of bias in results, rather lack of statement for COI which, if not mandatory by journal policy, is often forgot by the authors. We know that the highest risk of bias affects RCT rather than cohort studies, and major drive to publish and ideally find significant results in cohort studies is for scientific reasons (and, being cynical, to improve personal background), but results from cohort studies are hard to affect (I don't see researcher inventing results or subverting common methods to try to obtain favorable results). In conclusions, authors should relax the conclusions of both abstract and text, authors may say "we detected potential risk of bias etc.", but quantitative results were not significant so makes no sense saying "they are in that way DESPITE not significant" (after all, heterogeneity between studies is very low, so the example of whole grains does not fit optimally to search for potential influence of industry on results). Notes on lack of COI are ok. A final short note, I don't have personal conflict of interest in providing this revision, it is solely my opinion.

Response: We have amended the concluding remarks in both the abstract and the manuscript.

Abstract (line 50, p 2) "We did not establish that the presence of food industry sponsorship or authors with a COI with the food industry, was associated with results or conclusions that favour industry sponsors. The association of food industry sponsorship or authors with a COI with the food industry and favorable results or conclusions is uncertain. However, our analysis was hindered by the low level of COI disclosure in the included studies."

(line 476, p 24) "We did not establish that the presence of food industry sponsorship or authors with a COI with the food industry, was associated with results or conclusions that favour industry sponsors. The association of food industry sponsorship or authors with a COI with the food industry and

favorable results or conclusions is uncertain. However, our analysis was hindered by the low level of COI disclosure in the included studies"

Manuscript (line 447, p23) has been edited: "However, until the influence of industry sponsorship in primary nutrition studies has been further explored and measured with larger samples of industry studies or studies that have author disclosure statements, this bias may still be unaccounted for in dietary guidelines."

Reviewer: 3

Reviewer Name: Eva María Navarrete-Muñoz Institution and Country: CIBER of Epidemiology and Public Health, Spain Please state any competing interests or state 'None declared': None declared

The study addresses an important issue about the possible influences of the sponsorship in the studies published about the relationship between whole grain and cardiovascular disease. This is a well written manuscript and focused on an interesting topic for BMJ Open's readers. However, there are some imprecision and some concerns about the methodology used that should be clarified.

a) The most important concern is about the search strategy. It is should be more specific to respond of the aim of the study, for example, the search included a randomized control trial or intervention but the aim of the review should have only included observational studies.

Response: Thank you for raising this point. We did not include the RCTs and observational studies from our search in the same review because different methods are needed to assess their risks of bias and they cannot be quantitatively combined. However, it is more efficient to search for the RCTs and observational studies in the same search as the PICO questions were the same.

Our manuscript states, (line 151, p8) 'The search also included terms for randomized control trials to identify relevant trials for a future systematic review' and in our protocol (supplementary file 2), pp 49, point 22, line 43 'Types of study to be included Give details of the study designs to be included in the review. If there are no restrictions on the types of study design eligible for inclusion, this should be stated. Inclusion: RCT/ cluster RCT Controlled Trial/ pseudo-randomized Cohort Case-control Pre/Post'

Our eligibility criteria state (line 157, p8): "This review included primary nutrition studies of cohort or case control designs that quantitatively examined the benefits or harms of wholegrain consumption related to cardiovascular disease outcomes in healthy children and/or adults."

We have added a sentence (line 156, p8) clarifying that for this particular review: "The randomized controlled trials identified in our search were included in another review currently under development. We selected observational studies for this review."

b) This review was conducted including observational studies (see lines 130-131), thus why did the authors decide to search in Cochrane Library?

Response: The Cochrane Library includes observational studies and reviews that contain observational studies.

c) In the line 146, the authors said this review included healthy children but it is impossible they found children studies using outcome measurements defined in the lines 160-165.

Response: Thank you for this comment. It is feasible for observational studies to enroll healthy children as participants into a study and follow them up over a long enough period of time that would allow for these clinical outcome measures to be assessed. An example is the Framingham studies

which enroll healthy children. As our sample did not include any studies with children we have not elaborated on this point in the text.

d) The abstract is so confused and it should be rewritten. In the results, they showed critical risk of bias but this concept was not previously defined. The aim is more clearly described in the article than in the abstract.

Response: We have rewritten the abstract to include the wording for the objective used in the paper. Our methods section (line 274, p14) states “Each domain is assessed at a low, moderate, serious or critical risk of bias, or no information.” The number of studies with a serious or critical risk of bias in each bias domain is presented in Table 1 p 17-18, and the number of studies with an overall serious or critical risk of bias is presented in the text (line 340-341, p18) and Figure 2. Therefore, it is appropriate to present the data on serious or critical risk of bias in the abstract.

e) In the background, we added information of previous studies which evaluate the role of sponsorship in studies with food. I recommended to read this publication Bes-Rastrollo et al. Financial conflicts of interest and reporting bias regarding the association between sugar-sweetened beverages and weight gain: a systematic review of systematic reviews. PLoS Med. 2013 Dec;10(12):e1001578;

Response: We agree that this an important piece of research and it has been included in our previous systematic review and meta-analysis examining the association of industry sponsorship with outcomes of nutrition studies ‘Chartres N, Fabbri A, Bero LA. Association of industry sponsorship with outcomes of nutrition studies: A systematic review and meta-analysis. JAMA Intern Med 2016;176(12):1769-77.’ We feel it is appropriate to cite the systematic review in the background section (Ref 20) as it covers studies of a variety of nutrition topics, not just sugar-sweetened beverages and weight gain.

f) In lines 192-195, the authors said that effect size was defined as the risk ratio but in some cases they used hazard ratios to present the results. Moreover, in the case-control studies, the adequate association measure is the odds ratio not risk ratio. The authors should change this sentence.

Response: We have now amended this sentence to state (line 205, p10) “Effect size was defined as the risk ratio, hazard ratio or odds ratio of the association between whole grains and a clinical outcome of cardiovascular disease.”

g) The same changes of the previous point should be conducted in the lines 272-273.

Response: We calculated risk ratios and hazard ratios in the meta-analyses. We have amended the sentence on line 284, p14: “We calculated risk ratios or hazard ratios (and 95% confidence intervals) to quantify the association between food industry sponsorship and / or authors with COI with the food industry and favourable results, favourable conclusions and the overall study risk of bias rating.”

h) In the data collection and analysis, the author mentioned an intervention period although the review only included observational studies.

Response: Line 244, p12 we changed the word “intervention” to “exposure duration”

i) In line 283, the authors said that they undertook random effects meta-analysis to compare the association across subgroups. However, they should use random effects when I2>50% while fixed effects were observed in the rest of the cases.

Response: We chose a random effects model because we anticipated that effects would follow a distribution across studies. Most textbooks on meta-analysis discourage post hoc selection of an effect model based on heterogeneity statistics (e.g., Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. Chichester, UK: John Wiley and Sons 2011).

j) In the table 1, some categories should be reduced in order to avoid some collapse, especially in risk bias assessment.

Response: We think this data would be hard to interpret if we did not collapse some categories. The breakdown for each risk of bias assessment (low, moderate, serious or critical risk of bias, or no information) by domain is shown in Figure 2.

k) The results should be reorganized, firstly they should include the answer to primary aim and after the answer to secondary aim.

Response: Thank you for this comment, however we believe that we have put these results in the correct order according to our Primary outcomes - Statistical significance of results favourable to the sponsor (line 349-355, p18), Effect size of results (line 357-372, p19) and Conclusions (line 374-380 p19-20) - and Secondary Outcomes -The risk of bias of the included studies (line 382-386, p20). This is the same order as presented in the Methods section. Concordance between study results and conclusions was not measured and we comment in line 288, p14 under Analysis, that 'We had planned to calculate a RR for level of concordance, however since in all studies there was concordance between the results and conclusions, we did not undertake this analysis'.

The Results that are discussed before these primary and secondary outcomes (i.e Search results (line 311-313, p15), Characteristics of included Studies (line 315-330, p15-16) and Risk of bias in included studies (line 339-347, p18) we feel are appropriately placed.

Reviewer: 4

Reviewer Name: Jonathan R Treadwell

Institution and Country: ECRI Institute, USA Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below BMJ statistical review 11-2-18

Jonathan Treadwell, ECRI Institute

The association of industry sponsorship with outcomes of studies examining the effect of wholegrain foods on cardiovascular disease and mortality: Systematic review and Meta-analysis

To assist the BMJ, my comments are divided into Major and Minor

Major comments

Strict vs lenient usage of $p<0.05$. The authors suggest that in order for one of their included studies to legitimately conclude that there is an association between industry sponsorship and a positive effect of wholegrain foods, the study must have found $p<0.05$. Thus, they strictly used $p<0.05$ when examining the results of individual studies (with $p<0.05$ indicating favorable to industry, and $p>0.05$ indicating results unfavorable to industry). However, when describing their own results, they were less strict, and despite $p>0.05$ in several instances, allowed themselves a gray-area conclusion of "Studies with industry ties were more likely to have favourable results than those with no industry ties…although the result was not statistically significant". This choice in how to present their results ("more likely" despite $p>0.05$) makes the reader think they might have been biased to find the association for which they were looking.

A more even-handed wording would have been something like "Data were inconclusive with respect to the association between industry ties and favorable results, as the relative risk could be as high as

2.35 or as low as 0.88". So the authors appear to be guilty of the same "discordance" between results and conclusions that they had planned to analyzed in one of their secondary outcomes (if any studies had had such a discordance, which they did not).

Response: We agree with these comments. We feel that confidence intervals are more informative than p-values. This is why we emphasize the confidence intervals in the results, for example (line 350, p18): "The risk of reporting favourable outcomes was 44% higher in studies with industry sponsorship and/or authors with a COI with the food industry RR= 1.44 (95% CI 0.88-2.35). However, the confidence interval was wide and included differences in risks that were unimportant or operating in the opposite direction as plausible estimates."

We agree that, in an attempt not to repeat the results in the conclusions, the statistical significance was overemphasized. We have included the reviewer's suggestion and edited the first paragraph of the Discussion (line 389 p20) as follows:

"Observational studies examining the effect of wholegrain consumption on cardiovascular disease outcomes that were sponsored by the food industry and / or had authors with a COI with the food industry more often had favourable results than research not tied to the food industry. However, his finding was inconclusive with respect to the association between industry ties and favorable results, as the relative risk could be as high as 2.35 or as low as 0.88."

The authors might instead consider a Bayesian analysis of the results, and simply state that given the data, there is an X% chance (with X being perhaps 80%) that studies with industry ties are more likely to have favorable results. That way, your words can be faithful to the data, and you're not a slave to an arbitrary 0.05 cutoff, and also you avoid any impression of bias towards a positive association.

RESPONSE: We agree that a Bayesian analysis would make an additional interesting study. However, such an analysis would deviate from our pre-specified and published analysis plan. Furthermore, such a Bayesian analysis would make it difficult to compare our finding to the large body of evidence that has examined the association of other corporate sponsors with research outcomes (as cited in Discussion, lines 430, p 22)

BMJ reviewers should understand that the first analysis mentioned in the abstract, the one with RR=1.44, is based simply on 2x2 table where 8 of 9 industry-ties studies had a favorable finding, as compared to 8 of 13 non-industry-ties studies had a favorable finding. So it's just 89% vs 62%. This is a very low power statistical test, since each study is simply put into one of four boxes in the 2x2 table…"vote-counting"…this low power perhaps explains the wide confidence interval. If 89% vs 62% had been based on more studies, such as 44 instead of 22, the same percentage difference would have been classified as statistically significant (16/18 vs 16/26 yields p=0.045).

A better statistical approach, which the authors have also done, is to use each study's quantitative effect size, and see if the weighted average of industry-ties-studies results was any different from the weighted average of non-industry-ties-studies' results. This is called subgroup analysis. For this analysis, the authors also found statistical non-significance, but they chose to describe their results in a very different way as "we did not find an important difference…". This is puzzling, because in both cases, there is a nonsignificant effect size that is in the direction (at least) of suggesting that industry ties are more likely to find favorable results. So why the completely different interpretation? Using a Bayesian approach throughout (and associated Bayesian wording for conclusions such as there is an X% chance that a hypothesis is true) may resolve this rather glaring inconsistency.

RESPONSE: We have edited the abstract to emphasize the large confidence interval. We think our methods and results make it clear where the numbers in the abstract originate and all the numbers in the abstract are in the results. We think it is appropriate to include the findings from all 3 of our a priori pre-specified outcomes in the abstract, rather than just one outcome (effect size). These

outcomes are stated in the methods section, under the heading “Primary Outcomes:” (lines 190-214, p 10-11) statistical significance of results favourable to the sponsor, 2) effect size of results, and 3) conclusions.

We have edited the reporting of the result on effect size in the abstract (line 45, p2) to delete the word “important:” “We did not find a difference in effect size (magnitude of RRs) between studies with industry ties, RR = 0.77 (95% CI 0.58-1.01) and studies with no industry ties, RR = 0.85 (95% CI 0.73-1.00) (P=0.50) I2 0%.”

As noted above, we think a Bayesian analysis is an interesting idea for a future study.

I would say that for the subgroup analysis, rather than presenting separately each subgroup’s summary effect size, you should present the statistical contrast between subgroups (its point estimate and CI). This output is not in Figure 3 or 4, but the output does give the exact p value (“Test for subgroup difference chi2=0.46 p=0.50). Figure out the corresponding point estimates and their CI’s, and present that.

RESPONSE: The calculation of subgroup differences when comparing study level variables rather than within study contrasts is not recommended. The test of interaction P-Value should only be presented. See for example: Fisher DJ, Carpenter JR, Morris TP, et al. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? BMJ 2017;356

Regarding your wording in the abstract “These findings suggest, but do not establish…”. This is very problematic wording. At heart, by the conventional standard for making scientific claims i.e. p<0.05, your results are inconclusive, and “inconclusive” should be your first word describing them. Notice that you felt that for the subgroup analyses you could firmly say “there is no important difference”. Why couldn’t the pro-wheat industry apply the same wording to your RR=1.44 finding due to the wide CI (ie they could call it no important difference)? You need to find a way to be consistent when interpreting different outcomes. I didn’t see any clear delineation in your methods or results for, if confronted with p>0.05, you would conclude 1 “no important difference” vs 2 you would say it’s inconclusive vs 3 you would use some complicated wording such as “suggests but do not establish”.

Response: Thank you for this comment and we have now amended the Conclusion in the Abstract (line 50, p 2) and manuscript (line 476, p 24) - ‘We did not establish that the presence of food industry sponsorship or authors with a COI with the food industry was associated with more favorable results or conclusions that favours industry sponsors. The association of food industry sponsorship or authors with a COI with the food industry and favourable results or conclusions is uncertain.”

Minor comments

The title should say “industry ties” instead of “industry sponsorship”. This is because authors defined (and analyzed) industry ties as including EITHER industry sponsorship OR industry COI’s amongst authors.

Response: This has now been amended ‘The association of industry ties with outcomes of studies examining the effect of wholegrain foods on cardiovascular disease and mortality: Systematic review and Meta-analysis’

In the abstract, within the results, it’s confusing that you’ve flipped the meaning of being RR>1 vs RR<1. When you present RR=1.44, clearly you defined RR>1 as meaning that industry ties are associated with more favorable results, but later when you present RR=0.77 and RR=0.85, you are then defining RR>1 as meaning the opposite. Please make these consistent. I do think the latter set of analyses are much more important, as they take advantage of quantitative effect sizes, rather than just put each study into one of four boxes.

Response: For effect size, we are just stating the RR and not giving any indication of direction. It is true that an RR of 1 would indicate equivalent effect sizes, but we are just reporting the actual RR calculated.

Line 196 specify whether 0.05 had to be one tailed or two tailed.

Response: We have added two-tailed (line 198, p 10).

You might consider replacing “favorable” with “pro-industry” and “unfavorable” with “anti-industry”. This might be a bit more direct and easier for readers.

Response: We appreciate this comment, however as we and others have used this language in previous reviews assessing the influence of industry sponsorship on other areas of research, we have retained this wording, see for example: Lundh A, Lexchin J, Mintzes B, et al. Industry sponsorship and research outcome. Cochrane Database of Systematic Reviews 2017. In addition, among researchers in the field, the terms “pro” and “anti” are considered inflammatory as they cover attitudes that extend far beyond the data we have analyzed.

Line 253. “non-profit organizations” What if there’s a non-profit organization that is anti-wheat? Or a pharma company that is anti-wheat? It’s too simplistic to put all of these into your “industry” category. I had thought, when reading the intro/methods that all your “industry” were prescreened to be pro-wheat industry. So these needs to be more obvious to readers that you’ve combined industries in this way.

Response: The industry of concern for this study is the industry that would have a relevant conflict of interest – the food industry. “Wheat” industry is not a classifiable category as many types of food companies make grain products. As described in the methods (lines 262, p13), we have ‘industry’ sponsored studies defined as “those declaring any sponsorship from the food industry, including if the study received ‘mixed funding’ from the food industry, non-profit organizations or other industries (i.e. pharmaceutical).” As we identified only 5 industry funded studies, and each of these studies had mixed funding (ie, food industry along with non-profit organizations or other industries), we could not analyse these separately.

Hazard ratios are A LOT like relative risks. You might consider doing a larger meta-analysis that includes both (and so perhaps ALL studies could be in it, for a new Figure 3). After all you included RR’s and OR’s in the same meta-analysis, and those are technically different (albeit similar given the low event rates).

Response: It is not recommended to combine hazard ratios with relative risks or odds ratios in a meta-analysis because hazard ratios represent instantaneous risk over the study time period, whereas RRs and ORs estimate risk/odds at a fixed time point. The denominator for HRs includes time, while the denominator for RRs and ORs does not, so HR and RR or OR are estimating different things. See, for example: Tierney JF, Stewart LA, Ghersi D, et al. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8:16-16.

VERSION 2 – REVIEW

REVIEWER	Marion Nestle New York University I earn royalties from books and honoraria for lectures about matters relevant to this paper.
REVIEW RETURNED	21-Dec-2018

GENERAL COMMENTS	The authors have adequately addressed the concerns noted in my initial review.

REVIEWER	Giuseppe Grosso University of Catania, Italy
REVIEW RETURNED	29-Dec-2018

GENERAL COMMENTS	The article is now suitable to be published