thebmj

BMJ 2019;364:l341 doi: 10.1136/bmj.l341 (Published 7 February 2019)

ANALYSIS

PREDIMED trial of Mediterranean diet: retracted, republished, still trusted?

Arnav Agarwal and **John P A Ioannidis** consider what we can learn from the retraction and republication of an influential trial of Mediterranean diet

Arnav Agarwal *research associate*[1], John P A Ioannidis *professor*[2]

[1]Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [2]Stanford Prevention Research Center and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA 94305, USA; Correspondence to: J P A Ioannidis jioannid@stanford.edu

The Prevención con Dieta Mediterránea (PREDIMED) trial[1] is one of the most influential randomised trials ever. It was cited 3364 times in Google Scholar in the five years after its publication. However, in June 2018 the trial was retracted and republished because serious protocol deviations were detected. Moreover, the repercussions of these protocol deviations and of the correction process raise many additional important questions. How do you correct one of the most influential trials and the large universe of its secondary publications?

Initial results and early raised concerns

PREDIMED was originally published in 2013.[1] Heralded as a pioneer effort in nutrition,[2] it randomised 7447 participants to a Mediterranean diet supplemented with extra virgin olive oil, a Mediterranean diet supplemented with nuts, or a control diet. It showed a 30% relative risk reduction in a composite clinical endpoint of cardiovascular death, stroke, or myocardial infarction in the Mediterranean diet groups.[1] PREDIMED was an important effort and one of the few long term trials with clinical outcomes in nutrition.

However, some concerns were raised at the time of the publication. The interventions tested were not a typical Mediterranean diet but single food supplementations. The control group experience was not necessarily non-Mediterranean. The primary outcome was a composite of three endpoints,[3 4] and significant differences were driven by a single endpoint (stroke) without differences in other cardiovascular disease or death. Effect sizes were probably inflated because the trial was stopped early after interim analyses showed benefit. Several systematic reviews and guidelines have either omitted PREDIMED[5 6] or have rated it as having serious risk of bias and being difficult to interpret.[7] Moreover, secondary PREDIMED analyses reported results that were deemed implausible.[8 9]

Retraction and republication

Recent developments questioned PREDIMED at its core. An analysis of reported baseline characteristics in 5087 trials by Carlisle identified trials in which the compared randomised groups were too similar or too dissimilar, raising questions of potential fraud or non-random sampling.[10] PREDIMED stood out for implausible P value patterns when comparing the baseline characteristics of the three arms. An audit of the trial found serious irregularities: enrolment of household members without randomisation; assignment of participants to study arms based on clinic site rather than true randomisation; and inconsistent use of randomisation tables. These deviations affected 1588 participants (21% of the total).[11]

The randomised trial was no longer a randomised trial. The original paper was retracted and replaced with a reanalysis that treated PREDIMED as a non-randomised study and excluded participants who were not truly randomised.[11] The reanalyses gave similar point estimates for the primary endpoint.

Is republication justified?

Whether republication is justified in such cases is controversial. Retracting a paper and replacing it with a republication is an uncommon choice. In theory, it can be used when there is an error that significantly affects parts of the study but does not completely refute it. Therefore, the validity of this course of action depends on the nature of the errors and on whether there is really a way to fully correct them.[12]

In PREDIMED, the detected irregularities may not entirely explain the peculiar baseline characteristics and they also raise questions about the quality of other aspects of the conduct of the trial, such as data collection, data arbitration, and adjudication. Participants, investigators, and assessors were not blinded, and the unmasked design can further compound any bias—for example, if investigators and sponsors favour specific interventions or when assessors are collecting outcome data. Here, it is unclear if the correction of the specific identified errors also corrected all the potential accompanying problems and consequences of these errors.

Additionally, the title of the republication does not make it clear that it is a reanalysis and republication, and many readers may be confused. Showing clearly that the new paper is an amendment and numbering versions would help.[13]

Continuing follow-up

If randomisation problems had been detected while the study was ongoing, would it have been stopped early because of

perceived benefit? The reanalysis excluding improperly randomised patients does not satisfy the P value boundary required for early stopping for each intervention arm. PREDIMED has in fact continued follow-up, and the investigators have published papers with 1.2 additional years of follow-up, during which the number of participants experiencing an event included in the composite primary endpoint increased by 19%. However, information is not provided on the comparison of the three arms for the primary endpoint with this extended follow-up. The full follow-up data should be disclosed and analysed for an intention-to-treat comparison to determine whether there is still benefit.

Full follow-up would also allow more complete assessment of mortality differences. Mortality is linked to a national registry and thus cannot be biased from subjective interpretation by the local investigators who collated data on clinical events before sending them to an adjudication committee. In the original follow-up, the control arm had 114 deaths, versus 116 and 118 in the two experimental arms, respectively.[1] More deaths have occurred during the additional 1.2 years' follow-up. Analysis of the full mortality data would provide the most convincing evidence from this trial and is essential to understand the robustness of its conclusion.

We have asked the corresponding authors of PREDIMED papers and (on their recommendation) the head of the steering committee to provide the number of primary outcome events and deaths in each arm in the extended follow-up. PREDIMED investigators responded that "information you have asked for is the main topic of ongoing analyses on PREDIMED data, as already approved by the steering committee." We asked them again whether they could share minimal information for this article but have not heard back from them over the past six months.

What about secondary publications?

Importantly, the original PREDIMED paper had already generated 267 secondary publications before its retraction and republication.[14] Thirty two of them have already received over 100 citations each in Google Scholar. Most of the publications come from the network of investigators who performed the original trial and their extended teams, with three investigators having each published over 150 articles from PREDIMED. In July 2018 we identified 203 secondary papers with data (excluding reviews, editorials, and commentaries); 194/203 (95%) first authors, 201/203 (99%) last authors, and 223/225 (99%) corresponding authors (some papers have more than one corresponding author) belong to the original PREDIMED investigator team in the 2013 paper or are affiliated with related Spanish institutions (see supplementary data on bmj.com).

The analyses presented in secondary papers use the data that led to retraction of the original. These publications should probably have notices of concern (as has been done for a secondary analysis published in *CMAJ*[15]) until they are properly re-evaluated. The PREDIMED authors have started correcting some of their work and have published several letters to this effect (at least five letters pertaining to eight secondary publications[16-20]).

Re-evaluation should be truly independent. Given the circumstances, it cannot be done only by the PREDIMED investigators and other sympathetic investigators who hold similar views on the importance of specific foods to modulate disease risk and on the agreement between the results of non-randomised studies (what PREDIMED in perceived to be now) and randomised trials. Involvement of investigators with contrary views is pivotal for this reassessment to be fair and balanced. Even the best intentioned and most knowledgeable investigators in nutrition may still favour their beliefs and thus inadvertently introduce some bias in the reanalyses.[21]

Instead of trying to correct one paper at a time, it may be more efficient for an independent team to make a centralised effort and report its findings to all relevant journals. Audits of bodies of contested literature are not uncommon. Since the work usually pertains to single investigators they are normally done by university appointed committees. For a major multi-investigator effort like PREDIMED, the independent assessors should have international provenance.

Secondary publications that compare outcomes in the randomised arms are directly affected. However, even when secondary publications deal with the study dataset or subsets as an observational cohort, the clustering of recruited participants (eg, household members co-randomised or a whole village recruited in one step) still affects their results.[22] The clustering effect also needs to be incorporated in observational analyses and may lead to different estimates and conclusions.

Inconsistencies in secondary publications

Even without in-depth re-evaluation of the raw data underlying PREDIMED publications, there are some inconsistencies in the reported data in these papers that suggest broader and more generalised problems that are not fixable by a single reanalysis.

To illustrate this point, we searched PubMed using the keyword "PREDIMED" and identified English language PREDIMED publications that included data on over 7000 participants reporting either the primary composite endpoint or mortality (all causes and cause specific). The number of events varies widely even when publications have similar follow-up (see supplementary table on bmj.com).

Some discrepancies may simply reflect missing data and different eligibility criteria in different analyses. However, table 1 shows a sample of contradictions in the reported data in paired papers that cannot be explained in this way and point to errors in one or both of the papers. The original publication[1] reported 87 cardiovascular deaths for the full trial population and follow-up, while Henríquez-Sánchez and colleagues' secondary analysis of dietary antioxidants and mortality[24] reports 102 such deaths despite a more limited sample and follow-up. Two publications[23 24] have identical follow-up, but one[23] has a larger number of total deaths, while the other[24] has a higher number of cardiovascular deaths. Two other publications[25 26] have identical total number of deaths, but one has five more cardiovascular deaths (a component of the primary endpoint)[26] while the other has five more deaths from other causes (not a component of the primary endpoint).[25] The original publication reported only cardiovascular deaths and total deaths.[1] However, if the 166 deaths from cancer reported by Henríquez-Sánchez and colleagues[24] are added to the 117 deaths from non-cancer, non-cardiovascular causes reported by Martínez-González and colleagues[25] and 87 deaths from cardiovascular causes in the original paper,[1] the total deaths are 370, exceeding the total deaths (n=348) reported in the original publication despite identical length of follow-up.[1]

These discrepancies may point to poor reporting, erroneous or inconsistent statistical analyses, or deeper problems related to problematic data collection and curation. They are superimposed on a PREDIMED literature that shows all the hallmarks of data dredging, given the huge number of secondary publications. Most analyses are not prespecified or are specified imprecisely. In fact, it is unclear which (if any) secondary analyses were

clearly and unambiguously prespecified. For example, one secondary publication states that invasive breast cancer was a prespecified outcome,[27] but the published trial protocol states that all cancer—not just invasive breast cancer—was a secondary outcome. Data are reported on one cancer subtype (claiming huge benefits) but not on dozens of others.

Publicly available data and reanalysis needed

PREDIMED investigators have stated that the data that went into the recent republication can be requested by interested parties, but the concerns identified make a strong case for the complete PREDIMED dataset, not just the data that went into the recent republication, to become publicly available. A truly independent audit should examine the original data records, adjudication, and statistical analyses that underlie this voluminous published literature. Updated follow-up results should be independently assessed and reported, including all-cause mortality data (the most objective outcome). Box 1 summarises some proposed actions. Some may wonder whether it is worth investing so much effort auditing a single study when there is a vast literature on nutrition with potential errors and biases. However, PREDIMED is an iconic example of using randomisation in the field (despite all the caveats discussed above) and it has major repercussions. An independent reanalysis effort may be more efficient and convincing than piecemeal attempts by the authors to defend each secondary publication separately.

Box 1: Potentially useful actions for PREDIMED

- Disclose full long term outcomes with updated follow-up
- Issue notices of concern for secondary publications until they are reassessed
- Consider centralised effort to re-evaluate all PREDIMED publications together
- Include both sympathetic and contrarian researchers in the re-evaluation
- Correct obvious inconsistencies that already violate plain logic rules
- Audit raw data, not just the clean data
- Audit data collection and curation procedures
- Correct or retract secondary publications, as appropriate
- Make raw data widely available (not only those pertaining to recent reanalysis)

Lessons for future pivotal multicentre trials

PREDIMED may offer useful lessons about how to run future large multicentre trials that aim to revolutionise an entire field. The original PREDIMED publication[1] had 18 authors and 223 collaborators. Many secondary papers also feature impressive numbers of coauthors. However, studies with hundreds of investigators may still have blind spots where no one really is responsible or knowledgeable enough to avert major mistakes and protocol deviations (box 2).

Box 2: Potential safeguards for large scale multicentre clinical trials

- Include strong methodologists in the statistical team
- Include strong methodologists in the data and safety monitoring board (DSMB)
- Investigators responsible for randomisation, data curation, and analysis should be designated and have sufficient methodological expertise
- Ensure that several people can independently safeguard against each type of error
- Monitoring board should include people who do not support the study hypothesis

PREDIMED investigators have a unique opportunity to disclose details on how the process failed, so that other trialists may avoid similar problems. Strong methodological expertise at all pivotal steps and function, both site specific and centralised, is essential. Also the inclusion of people who do not support the study hypothesis in the monitoring board is key to ensure balance and avoid bias.

PREDIMED offers useful lessons that can be applied to many other trials.[28] The problems should not lead to a reduction in funding of diet related research to improve health. Conversely, the same or even higher funding should be diverted to well executed large trials in nutrition. This will require getting together people with different expertise and skill sets. We also have an opportunity to probe how large volumes of published literature can be reassessed and corrected, as appropriate, when serious problems are identified. Finally, we can ponder on what expertise and better safeguards are needed to run such important multicentre trials reliably in the future.

Key messages

- PREDIMED, a highly influential trial of nutrition, was recently retracted and republished after major protocol deviations were noted
- Republication may not solve multiple problems that remain, including the inappropriateness of stopping early given the revised results and the effects on over 200 secondary publications
- Multiple contradictions between data reported across PREDIMED publications suggest a more generic problem with the trial's quality.
- PREDIMED may provide useful lessons on how to reassess and correct large volumes of published literature and on what methodological safeguards are needed for pivotal multicentre trials

Contributors and sources: JPAI conceived the idea for this article and both authors elaborated on it. Both authors extracted and evaluated data, wrote and revised the paper, and approved the final version. JPAI is guarantor.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that the Meta-Research Innovation Center at Stanford (METRICS) has been funded by the Laura and John Arnold Foundation. JPAI's work is funded by an unrestricted gift from Sue and Bob O'Donnell. JPAI loves olive oil, nuts, and the Mediterranean diet.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Estruch R, Ros E, Salas-Salvadó J, etal. PREDIMED Study Investigators. Primary prevention of cardiovascular disease with a Mediterranean diet [retracted]. *N Engl J Med* 2013;368:1279-90. 10.1056/NEJMoa1200303 23432189
2 Guasch-Ferré M, Salas-Salvadó J, Ros E, etal. PREDIMED Investigators. The PREDIMED trial, Mediterranean diet and health outcomes: how strong is the evidence?*Nutr Metab Cardiovasc Dis* 2017;27:624-32. 10.1016/j.numecd.2017.05.004 28684083
3 Montori VM, Permanyer-Miralda G, Ferreira-González I, etal . Validity of composite end points in clinical trials. *BMJ* 2005;330:594-6. 10.1136/bmj.330.7491.594 15761002
4 Ferreira-González I, Busse JW, Heels-Ansdell D, etal . Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ* 2007;334:786. 10.1136/bmj.39136.682083.AE 17403713
5 Rees K, Hartley L, Flowers N, etal . "Mediterranean" dietary pattern for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev* 2013;(8):CD009825.23939686
6 Hooper L, Martin N, Abdelhamid A, Davey Smith G. Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database Syst Rev* 2015;(6):CD011737.26068959

7 National Clinical Guideline Centre (UK). *Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease. London.* National Institute for Health and Care Excellence, 2014.
8 Guasch-Ferré M, Bulló M, Martínez-González MÁ, etal. PREDIMED study group. Frequency of nut consumption and mortality risk in the PREDIMED nutrition intervention trial. *BMC Med* 2013;11:164. 10.1186/1741-7015-11-164 23866098
9 Ioannidis JP. Implausible results in human nutrition research. *BMJ* 2013;347:f6698. 10.1136/bmj.f6698 24231028
10 Carlisle JB. Data fabrication and other reasons for non-random sampling in 5087 randomised, controlled trials in anaesthetic and general medical journals. *Anaesthesia* 2017;72:944-52. 10.1111/anae.13938 28580651
11 Estruch R, Ros E, Salas-Salvadó J, etal. PREDIMED Study Investigators. Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts. *N Engl J Med* 2018;378:e34. 10.1056/NEJMoa1800389 29897866
12 Fanelli D, Ioannidis JPA, Goodman S. Improving the integrity of published science: An expanded taxonomy of retractions and corrections. *Eur J Clin Invest* 2018;48:e12898. 10.1111/eci.12898 29369337
13 Barbour G. Amending published articles: time to rethink retractions and corrections?[version 1; referees: 2 approved with reservations]. *F1000 Res* 2017;6:1960. 10.12688/f1000research.13060.1
14 Predimed. Publications. http://www.predimed.es/publications.html
15 Mediterranean diets and metabolic syndrome status in the PREDIMED randomized trial. *CMAJ* 2018;190:E808. . 10.1503/cmaj.180791 29970374
16 Martínez-González MA. Protocol deviations, reanalyses, and corrections to PREDIMED Trial derivative study on peripheral artery disease. *JAMA* 2018;320:2272. . 10.1001/jama.2018.16553 30398615
17 Martínez-González MA. Protocol deviations, reanalyses, and corrections to derivative studies of the PREDIMED Trial. *JAMA Intern Med* 2018;178:1730-1. 10.1001/jamainternmed.2018.6456 30398517
18 Salas-Salvadó J, Bulló M, Babio N, etal. PREDIMED Study Investigators. Erratum. Reduction in the incidence of type 2 diabetes with the Mediterranean Diet: Results of the PREDIMED-Reus nutrition intervention randomized trial. Diabetes Care 2011;34:14-19. *Diabetes Care* 2018;41:2259-60. 10.2337/dc18-er10 30104300
19 Díaz-López A, Babio N, Martínez-González MA, etal. PREDIMED Study Investigators. Erratum. Mediterranean Diet, retinopathy, nephropathy, and microvascular diabetes complications: a post hoc analysis of a randomized trial. Diabetes Care 2015;38:2134-2141. *Diabetes Care* 2018;41:2260-1. 10.2337/dc18-er10a 30104299
20 Martínez-González MÁ, Hershey MS, Zazpe I, Trichopoulou A. Correction: Martínez-González, M.A. et al. Transferability of the Mediterranean diet to non-Mediterranean countries. What is and what is not the Mediterranean diet. *Nutrients* 2017, *9*, 1226. *nutrients* 2018;10:E823. 10.3390/nu10070823 29949867
21 Ioannidis JPA, Trepanowski JF. Disclosures in nutrition research: why it is different. *JAMA* 2018;319:547-8. 10.1001/jama.2017.18571 29222543
22 Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. *Ann Intern Med* 2001;135:112-23. 10.7326/0003-4819-135-2-200107170-00012 11453711
23 Martínez-González MA, Zazpe I, Razquin C, etal. PREDIMED GROUP. Empirically-derived food patterns and the risk of total mortality and cardiovascular events in the PREDIMED study. *Clin Nutr* 2015;34:859-67. 10.1016/j.clnu.2014.09.006 25304294
24 Henríquez-Sánchez P, Sánchez-Villegas A, Ruano-Rodríguez C, etal . Dietary total antioxidant capacity and mortality in the PREDIMED study. *Eur J Nutr* 2016;55:227-36. 10.1007/s00394-015-0840-2 25663609
25 Martínez-González MA, Sánchez-Tainta A, Corella D, etal. PREDIMED Group. A provegetarian food pattern and reduction in total mortality in the Prevención con Dieta Mediterránea (PREDIMED) study. *Am J Clin Nutr* 2014;100(Suppl 1):320S-8S. 10.3945/ajcn.113.071431 24871477
26 Hernández-Alonso P, Salas-Salvadó J, Ruiz-Canela M, etal . High dietary protein intake is associated with an increased body weight and total death risk. *Clin Nutr* 2016;35:496-506. 10.1016/j.clnu.2015.03.016 25886710
27 Toledo E, Salas-Salvadó J, Donat-Vargas C, etal . Mediterranean diet and invasive breast cancer risk among women at high cardiovascular risk in the PREDIMED Trial: a randomized clinical trial. *JAMA Intern Med* 2015;175:1752-60. 10.1001/jamainternmed.2015.4838 26365989
28 Ebrahim S, Montoya L, Kamal El Din M, etal . Randomized trials are frequently fragmented in multiple secondary publications. *J Clin Epidemiol* 2016;79:130-9. 10.1016/j.jclinepi.2016.05.016 27387965

Table

Table 1| Examples of inconsistent data on main endpoints across some PREDIMED publications

Publication	Median follow-up (years)	No of participants*	Primary composite endpoint	Mortality			
				All causes	Cancer related	Cardiovascular	Other
Martinez-Gonzalez 2015[23]	4.3	7216	277	328	—	81	—
Henriquez-Sanchez 2016[24]	4.3	7015	—	319	166	102	—
Martinez-Gonzalez 2014[25]	4.8	7216	—	323	130	76	117
Hernandez-Alonso 2016[26]	4.8	7216	277	323	130	81	112
Estruch 2013[1] (original publication)	4.8	7447	288	348	—	87	—

* Differences in sample size across studies are mainly because of different exclusion criteria and may be justifiable. However, even then inconsistencies are noted. For example, the first three publications in the table[23-25] all report that they exclude 231 participants but Martinez-Gonzalez 2015[23] states that they all had extreme values of total energy intake, whereas the other two state that some had extreme values of total energy intake (n=153[25], n=152[24]) and other had incomplete dietary data at baseline (n=78[25], n=79[24]). Also these papers give different median follow-up despite the similar exclusions.