Research methods and reporting

Thresholds for interpreting the fragility index derived from sample of randomised controlled trials in cardiology: a meta-epidemiologic study

Mohammad Hassan Murad [1,2] Abdalla Kara Balla,[2] Muhammad Shahzeb Khan,[3] Asim Shaikh,[4] Samer Saadi,[2] Zhen Wang [2]

10.1136/bmjebm-2021-111858

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjebm-2021-111858).

[1]Division of Public Health, Infectious Diseases and Occupational Medicine, Mayo Clinic, Rochester, MN, USA
[2]Evidence-based Practice Center, Mayo Clinic, Rochester, Minnesota, USA
[3]Division of Cardiology, Duke University School of Medicine, Durham, North Carolina, USA
[4]Department of Internal Medicine, Dow University of Health Sciences, Karachi, Pakistan

Correspondence to:
Dr Mohammad Hassan Murad, Public Health, Infectious Diseases and Occupational Medicine, Mayo Clinic, Rochester, Minnesota, 55905, USA; murad.mohammad@mayo.edu

To cite: Murad MH, Kara Balla A, Khan MS, *et al*. *BMJ Evidence-Based Medicine* 2023;**28**:133–136.

Abstract

The fragility index (FI) was proposed as a simplified way to communicate robustness of statistically significant results and their susceptibility to a change of a handful number of events. While this index is intuitive, it is not anchored by a cut-off or a guide for interpretation. We identified cardiovascular trials published in six high impact journals from 2007 to 2021 (500 or more participants and a dichotomous statistically significant primary outcome). We estimated area under curve (AUC) to determine FI value that best predicts whether the treatment effect was precise, defined as adequately powered for a plausible relative risk reduction (RRR) of 25% or 30% or having a CI that is sufficiently narrow to exclude a risk reduction that is too small (close to the null, <0.05). The median FI of 201 included cardiovascular trials was 13 (range 1–172). FI exceeded the number of patients lost to follow-up in 46/201 (22.89%) trials. FI values of 19 and 22 predicted that trials would be precise (powered for RRR of 30% and 25%; respectively, combined with CI that excluded risk reduction <0.05). AUC for meeting these precision criteria was 0.90 (0.86–0.94). In conclusion, FI values that range 19–22 may meet various definitions of precision and can be used as a rule of thumb to suggest that a treatment effect is likely precise and less susceptible to random error. The number of patients lost to follow-up should be presented alongside FI to better illustrate fragility.

Background

Due to many limitations and common misinterpretations of the p value,[1] the fragility index (FI) has been suggested as an easier, more intuitive way to communicate results to clinicians and other stakeholders.[2] The FI is defined as the minimum number of patients whose status would have to change from a non-event to event to turn a statistically significant result to a non-significant result. Thus, a randomised controlled trial (RCT) with statistically significant results that has an FI of 1 would lose significance even if one patient had the opposite outcome. FI was not intended to replace the p value, CI or precision judgements. Rather, it is intended to be a simple intuitive way to communicate findings to clinicians or the public.

A previous study evaluated cardiovascular RCTs with sample sizes over 500 participants that had a statistically significant primary outcome and showed a median FI of 13 (IQR, 5–26).[3] While intuitively one can think of an RCT outcome with FI of 1 or 2 to be less reliable, that is, susceptible to random error and erroneous misclassification of outcomes, it is not clear how to interpret FI of 5 or 6, for example. Thus, the lack of established cut-off or guide to aid in the interpretation of FI adds to some previously described[4–6] interpretational challenges.

Furthermore, modern frameworks of rating the certainty of evidence such as Grading of Recommendations, Development, Assessment and Evaluation (GRADE)[7] do not depend on statistical significance or the resultant calculation of FI. GRADE suggests that even if an estimate was statistically significant, it will not be considered precise (ie, robust or less prone to chance) unless it was derived from a body of evidence with a sample size that is adequate to detect a plausible relative risk reduction (RRR). GRADE suggests using RRR of 25%–30% for this estimation.[8] In addition to sample size considerations, GRADE suggests that judgements about precision should also consider whether the CI did not overlap a decision-making threshold that is considered to be trivial or unimportant.[8] Therefore, if the upper boundary of a relative risk is very close to the null or crosses a decision-making threshold, the results may still be considered imprecise despite statistical significance.

Considering the lack of anchors for FI and the lack of clarity about the relationship between FI and precision, we aimed to empirically evaluate FI in cardiovascular RCTs and study the association with precision. To date, this has not been studied and precision cannot be deduced from FI. Providing clinicians and other stakeholders with FI values that are likely to be associated with precise and reliable estimates can help them make judgements about certainty and trustworthiness of estimates.

Methods

This meta-epidemiological study follows the reporting guidance for methodology research.[9] A reporting checklist is provided in the online supplemental appendix. This study is a previously published protocol.[3] Since publicly available data were used, institutional review board approval was not applicable.

Data sources

Journals were selected for the present study based on a combination of the following features: impact factor, readership, specialisation in publication of cardiovascular RCTs and global recognition for consistent publication of influential RCTs over the last several decades. *The New England Journal of Medicine*, *The Lancet* and *Journal of the American Medical Association* were selected for having the highest impact factors in general medicine, while *Journal of the American College of Cardiology*, *European Heart Journal* and *Circulation* were selected for having the highest impact factors in the field of cardiovascular medicine. The rationale for targeting randomised trials with a sample size >500 and published in these specific journals was that we aimed to evaluate robustness in trials that were more likely to impact practice. We updated a previously published[3] search strategy through 13 September 2021. Details of the search strategy are available in the online supplemental appendix.

Study eligibility and data extraction

All RCTs were assessed for inclusion from the three cardiovascular journals whereas RCTs from the three non-cardiovascular journals were screened for determination of possible cardiovascular nature (if the interventions or outcomes were described as cardiovascular, such as those in the disciplines of heart failure, interventional cardiology, preventive cardiology, electrophysiology, cardiac imaging or stroke). Additional inclusion criteria were: (1) phase 3 or 4 RCT; (2) sample size ≥500 patients (an arbitrary cut-off to identify larger RCTs that are more likely to impact practice); (3) parallel arm study design and (4) at least one statistically significant binary outcome. Data were extracted in a pre-designed form. Study selection and data abstraction were performed by one reviewer (AKB) and verified by a second reviewer (SS). Discrepancies were reconciled by a third reviewer (MHM). Data were extracted using pre-defined forms that were pilot tested and included trials first author, year of publication, journal, impact factor, number of centres, country, a 2×2 table for the main outcome, number of patients lost to follow-up, funding, intervention type and control type.

Outcome measures

We evaluated the FI value that best predicts a precise treatment effect. Results were reported as the FI cut-off values and associated sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve (AUC).

Precision thresholds and definitions

Two methods were used to define a precise treatment effect: (1) whether an RCT was adequately powered for a RRR (25% or 30%); (2) whether the CI of the treatment effect was sufficiently narrow to exclude a small or trivial risk reduction of 0.05. The RRR thresholds of 25%–30% were recommended by the GRADE Working Group.[8] Precision guidance published in 2011 stated that although determining a threshold for adequate power is a matter of judgement and can change based on context, RRR 25%–30% can be considered a moderate or plausible RRR for most interventions, and can be used to determine whether a body of evidence had adequate sample size, assuming a type 1 error of 0.05 and a type 2 error of 0.20.[8] The thresholds for the second precision criteria of a CI boundary of RR of 0.05 was arbitrary. For the purpose of this analysis, we considered an RRR of less than 0.05 to be small or trivial, although we acknowledge that in a certain context such risk reduction may be relevant to some stakeholders.

Statistical analysis

Data from each RCT were presented in a 2×2 contingency table. The FI was calculated as described by Walsh *et al*.[2] Events were added to the smaller event group and non-events were simultaneously subtracted, while maintaining a constant patient population. The Fisher exact test was then used to recalculate the two-sided p value, while iteratively adding events until the p value reached or exceeded 0.05. The number of additional events required to reach a p value of ≥0.05 was defined as the FI. To determine whether an RCT had 80% power to detect a statistically significant difference using a χ^2 test with two-sided significance level of 0.05, we calculated the baseline risk for the control group and assumed a moderate RRR of 25% or 30% for the treatment group. We constructed ROC curves to predict FI values using a non-parametric model proposed by Pepe.[10] The sensitivity and specificity and the corresponding FI cut-off values were estimated using the minimised distance between the selected point on ROC curve and the perfect sensitivity and specificity.[11] The nearest to (0,1) method was used to find the cutpoint on the ROC curve closest to (0,1) (ie, the perfect sensitivity and specificity). We compared FI between trials that had FI less than the number of patients lost to follow-up, compared with trials that did not, using Mann-Whitney U test. We used the 'fragility' package, 'roctab' command and 'cutpt' package as implemented in Stata V.17.0 (StataCorp).

Results

Description of randomised trials

Database search identified 1365 potential citations from which 78 were included and added to trials identified in a previous study.[3] Therefore, we finally included 201 cardiovascular RCTs. The process of study selection is depicted in the online supplemental figure 1 and the list of RCTs is provided in online supplemental appendix table, along with their raw data, effect size and FI. Most RCTs were multicentred (93.3%). More than half of the RCTs (59.2%) had an active comparator and (62.7%) evaluated pharmacological interventions. The mean sample size of an RCT was 5234 participants (IQR: 1046–7046). The FI ranged 1–172 and had a median of 13 (IQR: 5–28). Eighteen RCTs (9%) had FI of 1. The description of included RCTs is provided in table 1.

FI cut-offs

Table 2 summarises FI cut-offs with highest AUC to predict precision based on whether the information size was sufficient (ie, the study had adequate power for RRR of 25% and 30%) or if the CI did not overlap an arbitrary decision-making threshold of 0.05. FI of 12 predicted that the RCTs would be powered for RRR of 25% or 30%. FI of 9 predicted that the CI excludes a risk reduction <0.05. FI of 19 predicted that RCTs would be powered for RRR of 30% and that the CI excluded a small risk reduction <0.05. FI of 22 predicted that RCTs would be powered for RRR of 25% and that the CI excluded a small risk reduction <0.05. AUC for meeting both of these two precision criteria was 0.90 (0.86–0.94).

Table 1 Description of 201 cardiovascular randomised trials

	Mean (range) or percentage
Multiple centres	93.0%
Active comparator	59.2%
Type of intervention	
Pharmaceutical	62.7%
Surgical	18.9%
Imaging	1.0%
Surgical	0.5%
Other	16.9%
Funding	
For profit	68.7%
Government	24.9%
Other	6.4%
Sample size	5234 (500–50,156)
Follow-up (months)	22.9 (1–118)
Loss to follow-up (%)	2.0 (0–26)
Fragility index	24.6 (1–172)
Sample size powered for 25% RRR	8274.2 (96–142 644)
Sample size powered for 30% RRR	5706.7 (68–96 236)

RRR, relative risk reduction.

FI exceeded the number of patients lost to follow-up in 46/201 (22.89%) trials. FI in this subset of trials was 40.33 (range 3–172); which was significantly higher than FI in trials that had FI equal or less than the number of patients lost to follow-up (FI 19.89, range 1–120, p value for the difference between the two FIs was 0.001).

Discussion

Studies that demonstrate statistically significant results provide evidence that rejecting the null hypothesis is less likely to be due to chance.[12] However, when such studies are underpowered, the possibility of both, type 1 and type 2 errors increases, and such results are labelled as fragile. Therefore, the FI was proposed as an intuitive and easy way to communicate statistically significant results to clinicians and other stakeholders including perhaps patients. This index has no known anchors or values at which the results would be considered adequate or robust. We evaluated the FI of modern and likely influential cardiovascular RCTs that enrolled 500 or more participants, published in high impact journals and had a statistically significant primary outcome. We report several key findings in this analysis. First, the current study has identified that FI values of 19–22 have the highest AUC (best combination of sensitivity and specificity) to predict that the estimates were precise. For decision-making purposes, RCTs with FI lower than this range are highly susceptible to chance and their results should be interpreted with caution.

A second important finding of this study is that many RCTs had FI of 1 and over half of them may not meet such precision cut-offs (median FI was 13). This means that if very few patients were re-classified in terms of having an event, the outcome would become statistically insignificant. Thus, the treatment effect of many cardiovascular RCTs remains fragile and susceptible to random error, despite their statistical significance. This finding of common fragility in trials has been observed in various fields such as cardiology, rheumatology, anaesthesiology, ophthalmology, critical care, spine surgery and sport medicine.[3 13–18] Lastly, almost 1 in 4 trials had FI that exceeded the number of patients lost to follow-up. Results of such trials are even more fragile and less robust because the patients lost to follow-up could be the patients who would have had a different outcome and would change the statistical significance of the difference between study arms. This finding provides a rationale for presenting the number of patients lost to follow-up alongside FI.

The implications of these findings to clinical practice are important. A well-known example of an RCT with FI of 1 that changed clinical practice was the one by Poldermans *et al*; which misleadingly suggested that perioperative beta blockers given to patients undergoing non-cardiac surgery reduce mortality. These findings were subsequently discredited and the routine implementation of the intervention has likely caused harm to many patients.[19] Evidence derived from trials with statistically significant results that are fragile should be labelled imprecise and warrant lower certainty. Low certainty should not lead to strong recommendations and universal implementation. In addition, FI values should be presented with additional information such as the number of patients lost to follow-up, as well as event rates and CIs.

Limitations and strengths

It is important to recognise that the current study has evaluated FI only as an intuitive way to present information to evidence users. It could be also used as a teaching tool. However, FI is certainly not a formal way to make judgements about imprecision and has limitations.[20–22] Imprecision judgements should be made using an established and rigorous approach based on CI and sample size considerations using context specific thresholds.[8] The thresholds we studied were arbitrary and may change based on the importance or nature of the outcome. Lastly, we anticipate that RCTs in lower tier journals may even have lower FI values because they will likely have smaller sample size. Lastly, decision making should depend on the totality of evidence synthesised in a systematic review, not an individual study.[23] FI is merely a way to present the finding of a single statistically significant RCT in a simplified way. FI does not change the binary view of hypothesis testing, but it adds nuance and communicate additional information beyond the binary view. For example, instead of saying: 'the results are significant', the FI index will inform stakeholders that

Table 2 Fragility index values predicting precision

Precision criteria	FI cutpoint	Sensitivity (%)	Specificity (%)	AUC	95% lower limit	95% upper limit
CI boundary closest to null ≥0.05	≥9	69.59	67.92	0.80	0.73	0.86
Sample size powered to RRR 25%	≥12	77.65	63.79	0.75	0.68	0.82
Sample size powered to RRR 30%	≥12	75.00	70.97	0.75	0.69	0.82
CI boundary ≥5% and RRR 25%	≥22	81.48	83.67	0.90	0.86	0.94
CI boundary ≥5% and RRR 30%	≥19	81.43	82.44	0.90	0.86	0.94

AUC, area under the receiver operating characteristic curve; FI, fragility index; RRR, relative risk reduction.

'the results are significant, but they would lose significance if two patients had a different outcome'.

Conclusions

The findings of this study demonstrate that FI values in the range of 19–22 can be used to suggest that a treatment effect is likely to be precise and less likely to be susceptible to random error. Contemporary cardiovascular RCTs with 500 or more participants that have statistically significant results have a median FI of 13. Thus, approximately half of them do not meet this proposed range of values. The findings also provide a rationale for presenting the number of patients lost to follow-up alongside FI.

Contributors MHM, ZW and AKB conceived the idea. AKB, MSK, AS and SS selected studies and extracted data. ZW conducted the analysis. MHM is the guarantor of this work.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs

Mohammad Hassan Murad http://orcid.org/0000-0001-5502-5975

Zhen Wang http://orcid.org/0000-0002-9368-6149

References

1 **Amrhein V**, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature* 2019;**567**:305–7.
2 **Walsh M**, Srinathan SK, McAuley DF, *et al*. The statistical significance of randomized controlled trial results is frequently fragile: a case for a fragility index. *J Clin Epidemiol* 2014;**67**:622–8.
3 **Khan MS**, Ochani RK, Shaikh A, *et al*. Fragility index in cardiovascular randomized controlled trials. *Circ Cardiovasc Qual Outcomes* 2019;**12**:e005755.
4 **Potter GE**. Dismantling the fragility index: a demonstration of statistical Reasoning. *Stat Med* 2020;**39**:3720–31.
5 **Carter RE**, McKie PM, Storlie CB. The fragility index: a P-value in sheep's clothing? *Eur Heart J* 2017;**38**:346–8.
6 **Porco TC**, Lietman TM. A fragility index: handle with care. *Ophthalmology* 2018;**125**:649.
7 **Hultcrantz M**, Rind D, Akl EA, *et al*. The grade Working group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 2017;**87**:4–13.
8 **Guyatt GH**, Oxman AD, Kunz R, *et al*. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol* 2011;**64**:1283–93.
9 **Murad MH**, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med* 2017;**22**:139–42.
10 **Pepe MS**. *The statistical evaluation of medical tests for classification and prediction*. New York: Oxford University Press, 2003.
11 **Perkins NJ**, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol* 2006;**163**:670–5.
12 **Royall R**. The effect of sample size on the meaning of significance test. *Am Stat* 1986;**40**:313–5.
13 **Berti A**, Cornec D, Medina Inojosa JR, *et al*. Treatments for giant cell arteritis: meta-analysis and assessment of estimates reliability using the fragility index. *Semin Arthritis Rheum* 2018;**48**:77–82.
14 **Mazzinari G**, Ball L, Serpa Neto A, *et al*. The fragility of statistically significant findings in randomised controlled anaesthesiology trials: systematic review of the medical literature. *Br J Anaesth* 2018;**120**:935–41.
15 **Ridgeon EE**, Young PJ, Bellomo R, *et al*. The fragility index in multicenter randomized controlled critical care trials. *Crit Care Med* 2016;**44**:1278–84.
16 **Evaniew N**, Files C, Smith C, *et al*. The fragility of statistically significant findings from randomized trials in spine surgery: a systematic survey. *Spine J* 2015;**15**:2188–97.
17 **Khan M**, Evaniew N, Gichuru M, *et al*. The fragility of statistically significant findings from randomized trials in sports surgery: a systematic survey. *Am J Sports Med* 2017;**45**:2164–70.
18 **Shen C**, Shamsudeen I, Farrokhyar F, *et al*. Fragility of results in ophthalmology randomized controlled trials: a systematic review. *Ophthalmology* 2018;**125**:642–8.
19 **Bouri S**, Shun-Shin MJ, Cole GD, *et al*. Meta-Analysis of secure randomised controlled trials of β-blockade to prevent perioperative death in non-cardiac surgery. *Heart* 2014;**100**:456–64.
20 **Acuna SA**, Sue-Chue-Lam C, Dossa F. The fragility Index-P values Reimagined, flaws and all. *JAMA Surg* 2019;**154**:674.
21 **Tignanelli CJ**, Napolitano LM. The fragility Index-P values Reimagined, flaws and All-Reply. *JAMA Surg* 2019;**154**:674–5.
22 **Andrade C**. The use and limitations of the fragility index in the interpretation of clinical trial findings. *J Clin Psychiatry* 2020;**81**. doi:10.4088/JCP.20f13334. [Epub ahead of print: 24 03 2020].
23 **Murad MH**, Montori VM, Ioannidis JPA, *et al*. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *JAMA* 2014;**312**:171–9.

Appendix

Reporting Checklist

Item	Page number or section where item is satisfied
Title	The title Identifies the report as a meta-epidemiologic study
Abstract	The background, goals, data sources, and key findings are presented in a structured abstract
Rationale	Background section, page 4
Objectives	Background section, page 4
Protocol	Protocol is unregistered, page 5
Eligibility criteria	Page 6
Information sources	Pages 5-6
Search	Page 5
Study selection	Page 6
Data collection process	Page 6

Data items	Page 6-7
Risk of bias in individual studies	Not relevant
Summary measures	Page 7
Synthesis of results	Page 7
Study selection	Page 8 and appendix
Synthesis of results	Page 8
Additional analysis	Sensitivity analysis, page 8-9
Summary of evidence	Page 9-10
Limitations	Page 10
Conclusions	Page 10
Funding	Page 2

Search strategy

("Lancet (London, England)"[Journal] OR "The New England journal of medicine"[Journal]) OR "Journal of the American College of Cardiology"[Journal]) OR "Journal of the American Medical Association"[Journal] OR "Circulation"[Journal] OR "European heart journal"[Journal] AND (Randomized Controlled Trial[ptyp] AND ("2017/12/31"[PDAT]: "2021/9/13"[PDAT])).

Study Selection

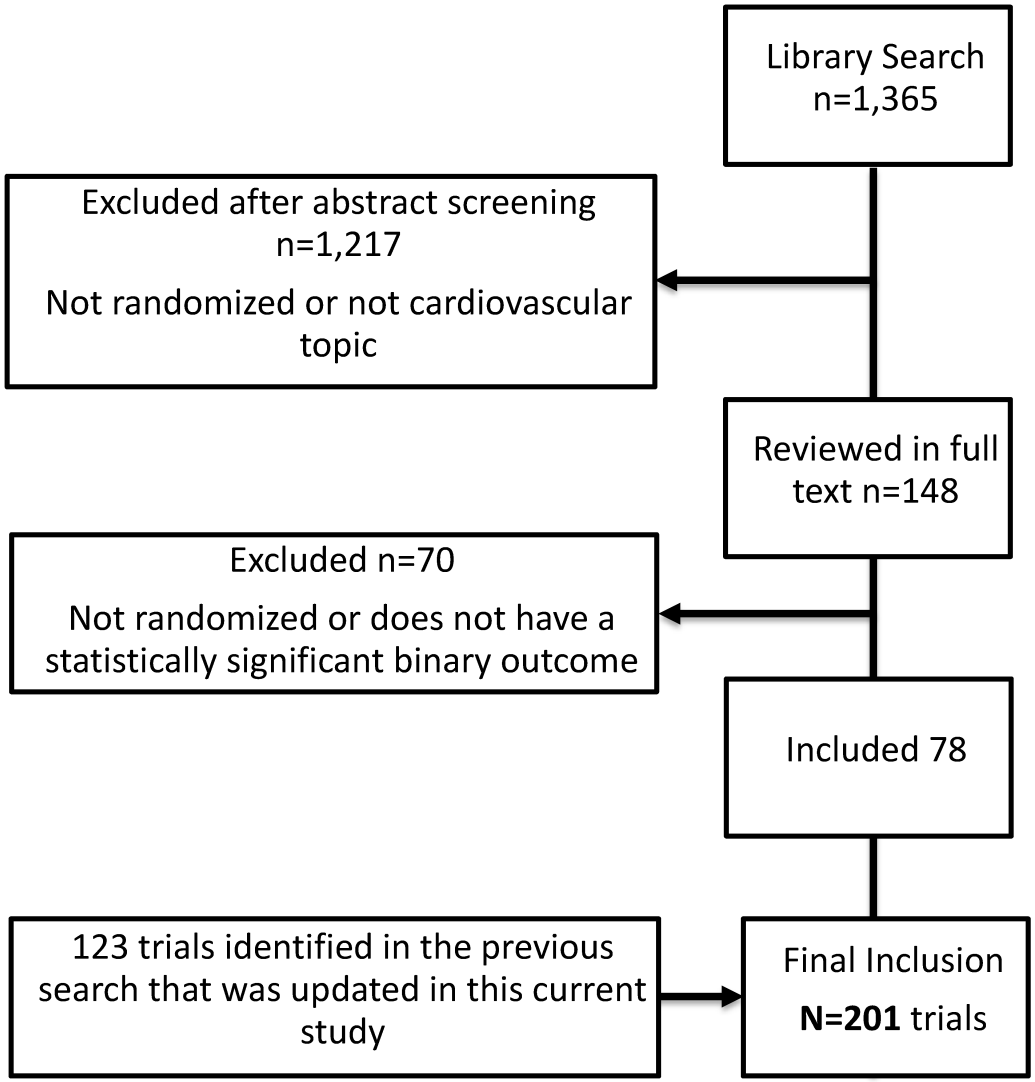

Table: Description of included trials

PMID	Year	Control Events	Control Sample	Intervention Events	Intervention Sample	RR	LL CI	UL CI	FI
34133859	2021	152	931	222	930	1.46	1.21	1.76	35
26551272	2021	354	4683	264	4678	0.75	0.64	0.87	41
33999547	2021	168	2391	114	2379	0.68	0.54	0.86	21
33459776	2021	369	1480	310	1523	0.82	0.71	0.93	24
33202219	2020	71	942	105	941	1.48	1.11	1.97	9
33200892	2020	355	614	245	608	0.70	0.62	0.78	72
33200891	2020	530	5292	400	5292	0.75	0.67	0.85	71
33197395	2020	294	550	217	558	0.73	0.64	0.83	49
33186492	2020	83	1421	59	1429	0.71	0.51	0.98	2
33185990	2020	1607	4112	1523	4120	0.95	0.90	1.00	1
32970396	2020	138	2152	100	2152	0.72	0.56	0.93	8
32882163	2020	116	1168	82	1170	0.71	0.54	0.93	8
32865380	2020	264	2760	187	2762	0.71	0.59	0.85	36
32865377	2020	462	1867	361	1863	0.78	0.69	0.88	50
32865376	2020	50	331	89	334	1.76	1.29	2.41	17
32865375	2020	316	6332	249	6399	0.78	0.66	0.92	24
32865374	2020	44	492	15	492	0.34	0.19	0.60	13
32668111	2020	362	5493	303	5523	0.83	0.72	0.97	11
32588060	2020	37	325	20	328	0.54	0.32	0.90	3
32434381	2020	20	1579	55	1572	2.76	1.66	4.59	17
32334703	2020	118	502	88	500	0.75	0.59	0.96	4
32272848	2020	17	421	140	487	7.12	4.38	11.58	80
32241367	2020	101	1678	32	1646	0.32	0.22	0.48	42
32227756	2020	61	389	75	388	1.23	0.91	1.68	1
32223113	2020	18	1795	4	1809	0.22	0.07	0.65	4
32222135	2020	584	3278	508	3286	0.87	0.78	0.97	18
32222134	2020	972	2524	897	2526	0.92	0.86	0.99	8

32112556	2020	96	713	39	527	0.55	0.39	0.78	12
31982074	2020	754	5764	668	5780	0.88	0.80	0.97	18
31779786	2020	83	676	57	663	0.70	0.51	0.96	2
31738483	2020	121	1430	156	1430	1.29	1.03	1.62	3
31733180	2019	38	818	64	826	1.67	1.13	2.46	6
31733140	2019	170	2379	131	2366	0.77	0.62	0.97	5
31577396	2019	17	297	29	284	1.78	1.00	3.17	1
31556978	2019	141	3555	250	3564	1.77	1.45	2.16	68
31553203	2019	21	400	45	400	2.14	1.30	3.53	8
31535829	2019	502	2371	386	2373	0.77	0.68	0.87	62
31479209	2019	156	1246	122	1242	0.78	0.63	0.98	3
31478763	2019	180	1642	123	1646	0.68	0.55	0.85	24
31475799	2019	184	2012	137	2006	0.75	0.60	0.92	12
31475798	2019	818	9601	736	9619	0.90	0.82	0.99	9
31475795	2019	339	2025	179	2016	0.53	0.45	0.63	113
31475793	2019	58	1108	35	1107	0.60	0.40	0.91	4
31448738	2019	301	3417	202	3421	0.67	0.56	0.80	55
31434507	2019	133	1695	89	1716	0.66	0.51	0.86	16
31216398	2019	209	464	162	454	0.79	0.68	0.93	13
31189511	2019	663	4952	594	4949	0.90	0.81	0.99	4
31185157	2019	30	1592	15	1591	0.50	0.27	0.93	1
31178152	2019	146	1894	61	1939	0.41	0.30	0.55	58
31141634	2019	85	429	117	439	1.35	1.05	1.72	5
30990260	2019	340	2199	245	2202	0.72	0.62	0.84	50
30883058	2019	68	454	42	496	0.57	0.39	0.81	11
30883052	2019	180	512	119	516	0.66	0.54	0.80	33
30786186	2019	27	421	11	420	0.41	0.21	0.81	3
30511879	2018	28	275	12	288	0.41	0.21	0.79	5
30415628	2018	606	4090	459	4089	0.76	0.68	0.85	85
30415601	2018	61	441	35	440	0.58	0.39	0.85	7
30403574	2018	1052	9462	903	9462	0.86	0.79	0.93	65

30291013	2018	428	4732	338	4731	0.79	0.69	0.91	36
30280640	2018	121	312	80	302	0.68	0.54	0.86	14
30261237	2018	39	724	21	724	0.54	0.32	0.91	2
30253879	2018	9	1245	1	1243	0.11	0.01	0.88	1
30221596	2018	265	9589	361	9525	1.37	1.17	1.60	49
30158069	2018	850	6276	1050	6270	1.24	1.14	1.34	120
30146931	2018	743	7740	658	7740	0.89	0.80	0.98	14
30145934	2018	81	2073	48	2073	0.59	0.42	0.84	10
30021076	2018	94	3995	130	4012	1.38	1.06	1.79	6
29900874	2018	133	877	97	877	0.73	0.57	0.93	7
29766772	2018	23	3604	62	3609	2.69	1.67	4.33	19
29766750	2018	160	2449	121	2432	0.76	0.61	0.96	6
29735587	2018	334	6214	266	6199	0.80	0.68	0.93	20
29544699	2018	10	1355	24	1357	2.40	1.15	4.99	2
29540324	2018	114	440	71	448	0.61	0.47	0.80	20
29527974	2018	199	3092	243	3098	1.22	1.02	1.46	4
29274727	2017	139	1540	305	1556	2.17	1.80	2.62	121
29231094	2017	21	524	36	522	1.72	1.02	2.91	1
29132879	2017	460	8261	347	8313	0.75	0.65	0.86	59
29102362	2017	98	604	68	596	0.70	0.53	0.94	5
17398308	2007	324	9319	262	9326	0.81	0.69	0.95	15
17448820	2007	121	669	68	666	0.56	0.43	0.74	26
17466227	2007	88	1218	65	1230	0.73	0.54	1.00	1
17693178	2007	48	485	24	488	0.50	0.31	0.80	8
17982182	2007	781	6795	643	6813	0.82	0.74	0.91	68
18071079	2008	41	1273	62	1280	1.50	1.02	2.21	1
18212283	2008	128	1035	95	1043	0.74	0.57	0.95	6
18280326	2008	32	301	13	299	0.41	0.22	0.76	5
18499566	2008	218	1802	166	1800	0.76	0.63	0.92	15
18757090	2008	765	3481	712	3494	0.93	0.85	1.02	1
18815396	2008	182	403	219	418	1.16	1.01	1.34	1

18997196	2008	251	8901	142	8901	0.57	0.46	0.69	67
19052124	2008	679	5762	552	5744	0.82	0.73	0.91	58
19144937	2009	91	496	67	509	0.72	0.54	0.96	4
19213680	2009	917	2327	734	2301	0.81	0.75	0.88	109
19228612	2009	159	903	105	897	0.66	0.53	0.84	23
19249633	2009	166	1765	115	1769	0.69	0.55	0.87	19
19336502	2009	924	3782	832	3772	0.90	0.83	0.98	17
19553646	2009	90	522	59	536	0.64	0.47	0.87	11
19664895	2009	52	330	32	338	0.60	0.40	0.91	4
19717846	2009	1014	9291	864	9333	0.85	0.78	0.92	73
19796737	2009	170	1153	93	2158	0.29	0.23	0.37	172
19812399	2009	39	453	68	445	1.77	1.22	2.57	11
19880844	2009	53	2026	101	2012	1.92	1.38	2.66	23
19922995	2009	889	1919	828	1927	0.93	0.86	1.00	4
20060578	2010	82	903	56	897	0.69	0.50	0.95	3
20079528	2010	668	6676	569	6732	0.84	0.76	0.94	38
20189239	2010	72	858	44	855	0.61	0.43	0.88	7
20194880	2010	20	250	38	250	1.90	1.14	3.17	3
20228403	2010	1722	4675	1532	4631	0.90	0.85	0.95	84
20382983	2010	26	602	11	614	0.41	0.21	0.83	3
20505177	2010	100	275	72	274	0.72	0.56	0.93	6
20691553	2010	189	245	211	256	1.07	0.98	1.17	1
20801500	2010	937	3264	793	3241	0.85	0.79	0.92	67
20818901	2010	490	4898	561	4906	1.14	1.02	1.28	10
21073363	2011	356	1373	249	1364	0.70	0.61	0.81	61
21073365	2010	364	904	297	894	0.83	0.73	0.93	23
21128814	2010	51	1705	36	1718	0.70	0.46	1.07	1
21175312	2010	74	2699	27	2708	0.36	0.23	0.56	26
21251705	2011	47	1201	75	1269	1.51	1.06	2.16	3
21315441	2011	120	280	84	270	0.73	0.58	0.91	10
21388309	2011	210	2215	178	2232	0.84	0.70	1.02	1

21870978	2011	265	9081	212	9120	0.80	0.67	0.95	12
22077816	2012	910	6471	822	6473	0.90	0.83	0.99	12
22082198	2011	19	1617	43	1619	2.26	1.32	3.86	7
22443427	2012	1176	13224	1028	13225	0.87	0.81	0.95	59
22453654	2012	74	356	51	356	0.69	0.50	0.95	3
22632908	2012	300	1520	245	1515	0.82	0.70	0.95	13
22858390	2012	105	501	68	500	0.65	0.49	0.86	12
22910755	2012	49	582	24	575	0.50	0.31	0.80	8
22924638	2012	56	441	19	447	0.33	0.20	0.55	20
23121323	2012	200	953	147	947	0.74	0.61	0.90	18
23265346	2013	40	282	15	250	0.42	0.24	0.75	7
23277305	2013	1225	9291	1057	9333	0.86	0.80	0.93	84
23415013	2013	126	279	54	284	0.42	0.32	0.55	52
23473369	2013	322	5470	257	5472	0.80	0.68	0.94	19
23484827	2013	90	242	64	233	0.74	0.57	0.96	3
23532240	2013	261	869	222	839	0.88	0.76	1.03	1
23652522	2013	557	954	346	948	0.63	0.57	0.69	165
23727163	2013	174	1437	122	1438	0.70	0.56	0.87	20
23808982	2013	49	2635	15	2609	0.31	0.17	0.55	16
23821090	2013	59	677	37	697	0.61	0.41	0.91	5
23995608	2013	80	5466	159	5106	2.13	1.63	2.78	56
24076283	2014	38	272	17	271	0.45	0.26	0.78	6
24076297	2013	58	1498	34	1500	0.59	0.39	0.89	5
24170388	2014	299	2660	237	2697	0.78	0.66	0.92	22
24171490	2013	94	1109	55	1089	0.60	0.43	0.82	14
24211309	2014	26	359	5	348	0.20	0.08	0.51	9
24678937	2014	6	357	23	390	3.51	1.45	8.52	5
24716680	2014	245	1723	206	1722	0.84	0.71	1.00	1
24716681	2014	28	499	13	506	0.46	0.24	0.87	3
24744272	2014	39	790	2	794	0.05	0.01	0.21	22
24780614	2014	149	506	92	506	0.62	0.49	0.78	29

24886787	2014	304	1460	257	1472	0.84	0.72	0.97	8
24963566	2014	9	285	45	286	4.98	2.48	10.00	20
25002178	2014	46	915	39	914	0.85	0.56	1.29	1
25085960	2014	83	903	36	897	0.44	0.30	0.64	24
25176015	2014	1117	4212	914	4187	0.82	0.76	0.89	118
25176289	2014	72	441	18	447	0.25	0.15	0.41	34
25399658	2014	65	4941	19	5020	0.29	0.17	0.48	26
25403646	2015	520	4187	604	4212	1.15	1.04	1.29	19
25465417	2014	851	10090	688	10080	0.81	0.73	0.89	87
25720624	2015	178	804	140	802	0.79	0.65	0.96	6
25771069	2015	355	10354	282	10348	0.79	0.68	0.93	23
25774645	2015	47	442	67	438	1.44	1.02	2.04	1
25791214	2015	429	4207	369	4197	0.86	0.76	0.98	7
25957224	2015	362	2585	275	2585	0.76	0.66	0.88	39
25980660	2015	91	416	87	414	0.96	0.74	1.25	1
26039521	2015	2742	9077	2572	9067	0.94	0.90	0.98	47
26061836	2015	172	360	188	305	1.29	1.12	1.48	23
26211828	2015	58	137	102	147	1.64	1.31	2.05	22
26267623	2015	17	5532	6	5576	0.35	0.14	0.89	2
26347918	2015	68	313	40	314	0.59	0.41	0.84	9
26408273	2015	30	308	15	307	0.50	0.28	0.91	2
26417061	2016	490	1022	417	1016	0.86	0.78	0.94	26
26466021	2015	113	1211	59	1221	0.52	0.38	0.70	28
26466202	2015	27	916	51	914	1.89	1.20	2.99	7
26551272	2015	319	4683	243	4678	0.76	0.65	0.90	29
26556051	2015	39	700	19	700	0.49	0.28	0.83	4
26699168	2015	242	582	155	574	0.65	0.55	0.77	51
26796390	2016	36	264	61	263	1.70	1.17	2.47	8
26886418	2016	228	1937	175	1939	0.77	0.64	0.92	16
27039945	2016	157	3168	113	3180	0.72	0.57	0.91	12
27040132	2016	304	6344	235	6361	0.77	0.65	0.91	24

27040723	2016	398	602	359	610	0.89	0.82	0.97	12
27195814	2016	148	1319	102	1317	0.69	0.54	0.88	15
27295427	2016	694	4672	608	4668	0.88	0.79	0.97	19
27539168	2016	15	404	38	401	2.55	1.43	4.57	9
27569841	2016	418	842	360	842	0.86	0.78	0.96	17
27810312	2016	121	592	81	592	0.67	0.52	0.87	14
27881569	2017	30	2991	14	2986	0.47	0.25	0.88	2
28304224	2017	1563	13780	1344	13784	0.86	0.80	0.92	118
28317415	2017	22	318	5	317	0.23	0.09	0.59	6
28510646	2017	85	323	43	322	0.51	0.36	0.71	21
28813218	2017	168	1104	131	1099	0.78	0.63	0.97	5
28844193	2017	264	981	151	981	0.57	0.48	0.68	75
28847206	2017	1803	15224	1640	15225	0.91	0.85	0.97	54
28851729	2017	5	501	19	500	3.81	1.43	10.12	4
28859942	2017	29	1094	139	1187	4.42	2.99	6.54	72
28859943	2017	2715	25078	2566	25078	0.95	0.90	0.99	14
28902590	2017	23	481	10	499	0.42	0.20	0.87	2
29083953	2017	158	344	189	341	1.21	1.04	1.40	7
29102362	2017	98	604	68	596	0.70	0.53	0.94	5

RR: relative risk. LL CI and UL CI are the lower and upper limits of 95% confidence interval.