Fragility index in fracture randomised controlled trials

Nick Tran, Thach S Tran, Tuan V Nguyen

Background

The traditional measures, such as relative risk, absolute risk and number needed to treat, are used to quantify the magnitude of effect size for an intervention (~ practical significance), and a p-value is used to assess whether the findings are due to chance (~ statistical significance). The p-value is theoretically the probability of the findings if the intervention is ineffective. Unfortunately, these metrics are not always intuitive to clinicians and are subject to misinterpretation, especially when event rates and sample sizes are small and follow-up times are limited[1,2].

Fragility index (FI) is a more intuitive metric, providing additional value over commonly reported p-value, risk reductions and confidence intervals[3-6] and facilitating the interpretation of the findings. FI is the minimum number of participants who must be moved from the non-event group to the event group to turn a statistically significant result into a nonsignificant one. Recently, some authors have also suggested the fragility quotient (FQ), i.e., the fraction of FI over sample size[7] as an extension of FI for measuring fragility of the significant dichotomous findings. FI has been widely used in other specialities[9-18], though no studies to date have been conducted to examine the fragility of findings in anti-fracture efficacy. We therefore conducted this systematic review to quantify the fragility of the positive findings in fracture prevention research using FI. The findings would help scientists make more informed decisions about the confidence warranted by fracture prevention results.

Study design

A systematic review is conducted to include randomised controlled trials (RCT) published within the last 30 years that provided statistically significant findings on the efficacy of pharmaceutical interventions in fracture prevention.

We used PubMed to systematically identify RCTs published in *(i) N Eng J Med, (ii) Lancet, (iii) BMJ, (iv) JAMA, (v) JAMA Intern Med, (vi) Ann Intern Med, (vii) PLos Medicine, (viii) J Bone Miner Res, (ix) J Clin Endocrinol Metab, (x) Bone,* and *(xi) Osteoporos Int.* The keywords include *"Bone Density Conservation Agents/therapeutic use"[MAJR] AND "Fractures, Bone/prevention and control"[MAJR] AND "Humans"[MeSH Terms] AND (randomized controlled trial [Publication Type] OR (randomized [Title/Abstract] AND controlled [Title/Abstract] AND trial [Title/Abstract])) AND (1992/07/31 [PDAT]: 2022/07/31 [PDAT]) AND ("N Engl J Med"[Journal] OR "Lancet"[Journal] OR "BMJ"[Journal] OR "JAMA Intern Med"[Journal] OR "Ann Intern Med"[Journal] OR "PloS Med"[Journal] OR "J Bone Miner Res"[Journal] OR "J Clin Endocrinol Metab"[Journal] OR "Bone"[Journal] OR "Osteoporos Int"[Journal]).* We will also search the reference lists of retrieved studies for additional studies.

We will include RCTs that (1) were two parallel involving humans, (2) allocated participants in a 1:1 ratio to intervention and control, and (3) reported, in the abstract at least one dichotomous or time-to-event outcome as statistically significant (P< 0.05 or a 95% CI excluding 1).

Statistical analysis

We first re-calculate a relative risk for all findings, including those with the time-to-fracture outcomes and exclude the findings with the re-calculated P-value> 0.05.

The primary analysis will include all significant findings from all primary, sensitivity, and subgroup analyses from the original studies or their extended components. The FI, FQ and corresponding IQR will be calculated for each finding using the "fragility" R package[19]. The fragility of anti-fracture efficacy will be described for all findings and doe specific subgroups using the median FI (IQR). We also report the proportion of findings with the reported number of participants lost to follow-up being greater than the responding FI.

The planned subgroup analyses include:

(1) Fracture sites, reported by the trial authors, as: (i) any fracture, (ii) osteoporotic fractures, (iii) major osteoporotic fractures, (iv) non-vertebral fractures, (v) vertebral fractures, (vi) clinical vertebral fractures, (vii) hip fractures, (viii) other fractures.
(2) Pharmacological intervention: (i) Bisphosphonates, (ii) Denosumab, (iii) Romosozumab, (iv) Teriparatide, (v) Strontium ranelate, (vi) Calcium with/without vitamin D.
(3) Fracture assessment timing: by 6 monthly.
(4) Sex: (i) Women, (ii) Men, (iii) Both.
(5) Nature of control group: (i) Placebo, (ii) Active placebo.
(6) Journals: (i) N Eng J Med, (ii) Lancet, (iii) BMJ, (iv) JAMA, (v) JAMA Intern Med, (vi) Ann Intern Med, (vii) PLos Medicine, (viii) J Bone Miner Res, (ix) J Clin Endocrinol Metab, (x) Bone, and (xi) Osteoporos Int.

We plan three sensitivity analyses:

(1) The analysis that includes only findings in which fracture was predefined as the primary endpoint of interest.
(2) The analysis that includes only highly statistically significant findings (i.e., P-value < 0.001).
(3) The analysis that includes only findings from the original studies (i.e., the findings from sensitivity or subgroups analyses or those from extended studies) will be excluded.

Dummy tables

Figure 1. Flowchart of number of papers screened and included in the analysis

Figure 2. Correlation between fragility index and sample size

Table 1. Fragility of anti-fracture efficacy

Table 2. Fragility of anti-fracture efficacy for findings with fracture as the primary endpoint

Table 3. Fragility of anti-fracture efficacy for findings with P< 0.001

Table 4. Fragility of anti-fracture efficacy for findings from the original studies

References

1. Breau RH, Dahm P, Fergusson DA, Hatala R. Understanding results. *J Urol* 2009;181:985–992. doi: 10.1016/j.juro.2008.11.029
2. Scales CD Jr, Norris RD, Preminger GM, Vieweg J, Peterson BL, Dahm P. Evaluating the evidence: statistical methods in randomized controlled trials in the urological literature. *J Urol* 2008;180:1463–1467. doi: 10.1016/j.juro.2008.06.026
3. Narayan VM, Gandhi S, Chrouser K, Evaniew N, Dahm P. The fragility of statistically significant findings from randomised controlled trials in the urological literature. *BJU Int*. 2018;122:160–166. doi: 10.1111/bju.14210
4. Akl EA, Briel M, You JJ, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. *BMJ* 2012;344:e2809. doi: 10.1136/bmj.e2809
5. Thorlund K, Imberger G, Walsh M, et al. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis–a simulation study. *PLoS One*. 2011;6:e25491. doi: 10.1371/journal.pone.0025491
6. 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;67:622–628. doi: 10.1016/j.jclinepi.2013.10.019.
7. Tignanelli CJ, Napolitano LM. The fragility index in randomized clinical trials as a means of optimizing patient care. *JAMA Surgery* 2019;154:74–79. doi: 10.1001/jamasurg.2018.4318, PMID: 30422256
8. Khan MS, Ochani RK, Shaikh A, et al. Fragility index in cardiovascular randomized controlled trials. *Circ Cardiovasc Qual Outcomes* 2019;12:e005755. doi: 10.1161/CIRCOUTCOMES.119.005755
9. Berti A, Cornec D, Medina Inojosa JR, Matteson EL, Murad MH. Treatments for giant cell arteritis: meta-analysis and assessment of estimates reliability using the fragility index. *Semin Arthritis Rheum*. 2018;48:77–82. doi: 10.1016/j.semarthrit.2017.12.009
10. Ridgeon EE, Young PJ, Bellomo R, Mucchetti M, Lembo R, Landoni G. The fragility index in multicenter randomized controlled critical care trials. *Crit Care Med*. 2016;44:1278–1284. doi: 10.1097/CCM.0000000000001670
11. 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;15:2188–2197. doi: 10.1016/j.spinee.2015.06.004
12. 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;45:2164–2170.
13. Shen Y, Cheng X, Zhang W. The fragility of randomized controlled trials in intracranial hemorrhage. *Neurosurg Rev*. 2019;42:9–14. doi: 10.1007/s10143-017-0870-8
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;120:935–941. doi: 10.1016/j.bja.2018.01.012
15. Matics TJ, Khan N, Jani P, Kane JM. The fragility index in a cohort of pediatric randomized controlled trials. *J Clin Med*. 2017;6:79
16. Checketts JX, Scott JT, Meyer C, Horn J, Jones J, Vassar M. The robustness of trials that guide evidence-based orthopaedic surgery. *J Bone Joint Surg Am*. 2018;100:e85. doi: 10.2106/JBJS.17.01039
17. Brown J, Lane A, Cooper C, Vassar M. The results of randomized controlled trials in emergency medicine are frequently fragile. *Ann Emerg Med*. 2019;73:565–576. doi: 10.1016/j.annemergmed.2018.10.037
18. Lin L, Chu H. Assessing and visualizing fragility of clinical results with binary outcomes in R using the fragility package. PLoS ONE 17(6): e0268754. doi: 10.1371/journal.pone.0268754

Table 1. Fragility of anti-fracture efficacy

	Subgroup	Number of trials	Fragility index (median [IQR])	Fragility quotient (median [IQR])
Overall				
Fracture sites	Any			
	Osteoporotic			
	Major osteoporotic			
	Non-vertebral			
	Vertebral			
	Clinical vertebral			
	Hip			
	Others			
Interventions	Bisphosphonates			
	Denosumab			
	Romosozumab			
	PTH analog			
	Strontium ranelate			
	Calcium/VitD			
Fracture assessment timing	0-6 months			
	6- 12 months			
	12- 18 months			
	18- 24 months			
	24- 36 months			
	36- 48 months			
	48- 60 months			
	>60 months			
Sex	Women			
	Men			
	Both			
Journal	NEJM			
	Lancet			
	BMJ			
	JAMA			
	JAMA Intern Med			
	Ann Intern Med			
	PLos Med			
	JBMR			
	JCEM			
	OI/Bone			
Control	Placebo			
	Active			