Fragility Index in Randomized Controlled Trials of Ischemic Stroke

Kenichiro Sato, MD, Tatsushi Toda, MD, PhD, and Atsushi Iwata, MD, PhD

Objective: The fragility index (FI), a minimum number of events in 1 arm of a clinical trial required to revert the statistically significant result to nonsignificant, has recently been developed as an easy-to-understand novel metric to evaluate the robustness of randomized controlled trials (RCTs). Here, we evaluated the FI of RCTs in the field of neurology, particularly in studies of ischemic stroke. *Methods:* Previous literature published between June 1, 2012 and May 31, 2018 were reviewed from the MEDLINE database by the authors. The original article reporting the significant RCT result, of which a dichotomous outcome was set as its primary outcome measure, was included to evaluate the robustness of the result by calculating the FI. In addition, recent studies examining FI in other clinical fields were reviewed and summarized. *Results:* In the 25 eligible RCT studies, the median total number of study participants was 206 (inter quartile range: 144-450) and the median FI was 7 (inter quartile range: 4-15.0). The FI showed a strong negative correlation with the observed *P* value. There was no significant difference in the FI between RCTs with and without acute settings. Our median FI was higher than the median FI of 2.5 of previous studies examining FI in other clinical fields, as only 20% (5 of 25) of studies included in our study had an FI less than 2.5. *Conclusion:* Our results suggest that many RCTs in the field of ischemic stroke have a fair robustness, when compared to those in other clinical fields.

Key Words: Fragility index—statistical robustness—RCT—ischemic stroke

Introduction

The fragility index (FI) is a recently developed measure of the minimum number of events (or nonevents) in 1 arm of a randomized controlled trial (RCT) required to change statistically significant results to nonsignificant.[1] A lower FI indicates that the significant study result is more likely to be altered to nonsignificant, meaning that the study has poor statistical robustness. The FI is an easy-to-understand metric to evaluate RCT results with a dichotomous endpoint, in addition to the conventional *P* value.

Since its introduction in 2014, as many as 10 studies have examined the FI in RCTs of the following clinical fields: spine surgery,[2] critical care,[3] pediatric medicine,[4] sports surgery,[5] intracerebral hemorrhage,[6] cardiology,[7] ophthalmology,[8] anesthesiology,[9] or head and neck surgery.[10] These studies often report that some RCTs in their field have a low FI.

Here, we applied this new metric to evaluate its robustness for RCTs in the field of ischemic stroke, which has thus far not been evaluated. In addition, we compared our results to FI studies from other clinical fields.

Methods

RCT Literature Review

Literature screening was performed by the author (K.S) on June 2018 using the MEDLINE database according to the abstract selection criteria. The PubMed query was as follows: (Search ((stroke[Title/Abstract]) AND randomized[Title/Abstract]) AND controlled[Title/Abstract] Filters: Clinical Trial; Randomized Controlled Trial; Abstract; Publication date from 2012/06/01 to 2018/05/31). The inclusion criteria were as follows:

From the Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.

Received December 18, 2018; accepted January 19, 2019.

Financial Disclosure: This study was supported by AMED under grant number 18dk0207020.

Address correspondence to: Atsushi Iwata MD, PhD, Department of Neurology, Graduate School of Medicine, The University of Tokyo, 113-8655, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. E-mail: iwata@m.u-tokyo.ac.jp.

1052-3057/$ - see front matter

https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.01.015

(1) original article reporting the results of a RCT for the treatment of ischemic stroke, (2) the results of the primary endpoint are significant with regard to their provided significance level, (3) the significant primary endpoint is a dichotomous variable, and (4) the article was published between June 1, 2012 and May 31, 2018, and the abstract is available in English. The selected articles were further reviewed, and all eligible studies were included in the analysis. If the goal of the RCT was to study an intervention for the chronic phase of acute ischemic stroke, or for the secondary prevention of ischemic stroke, the study was regarded as "not acute-setting." From the reviewed articles, we obtained the number of participants within each control and intervention subgroup, the number of events and nonevents, and the original level of significance. The FI in each study was calculated in accordance with earlier literature.[1-10] We have not obtained the number of "loss to follow-up" participants, as its definition and availability in each study were not always consistent.

In addition, recent studies examining the FI in other clinical fields were reviewed, and their FI (median and inter quartile range [IQR]) and number of participants (median and IQR for both events and nonevents) were summarized.

Statistical Analysis

All statistical analyses were performed using R (version 3.3.3). When comparing data between subgroups, we used the Fisher exact test for categorical data and the Wilcoxon rank sum test for continuous data. The differences were considered significant when the *P* value was <.05, unless otherwise mentioned.

Ethics

Ethical approval was not required for this type of study.

Results

From 1496 articles found in the PubMed/MEDLINE search, 25 eligible RCT studies were included in the analysis,[11-35] as summarized in Table 1. Overall, the median FI was 7 (IQR: 4-15.0), the mean FI was 10.9, and the studies with an FI less than or equal to 2 (lower FI) comprised 20% of the total (5 of 25). The FI distribution is plotted on the histogram shown in Figure 1. The FI had a strong negative correlation with the p value (rho =- .824, $P < .001$, Spearman's rank correlation; Fig 2, *P* value is shown with logarithm). The median total number of study participants was 206 (IQR: 144-450), and the

Table 1. *Eligible 25 RCT studies on ischemic stroke*

Ref#	Year	Author	Total event	Total sample size	Fisher_*P*	FI	Acute
11	2012	Rosso	143	176	<.001	17	Yes
12	2013	Asteriou	25	200	.009	3	No
13	2013	Cao	9	310	.03	2	No
14	2013	Higgins	24	100	<.001	10	No
15	2014	Reddy	73	707	.027	5	No
16	2014	Zhang	16	30	<.001	7	No
17	2014	Biase	41	1584	<.001	22	No
18	2014	Piironen	15	36	<.001	9	Yes
19	2015	Suntrup	17	30	.007	2	No
20	2015	Zaidat	19	111	.046	1	No
21	2015	Kuliha	56	150	.002	7	No
22	2015	Gong	61	450	<.001	12	No
23	2015	Huo	637	20693	.003	24	No
24	2015	Jovin	74	206	.029	2	Yes
25	2015	Saver	92	191	.001	10	Yes
26	2015	Goyal	130	311	.001	18	Yes
27	2015	Campbell	41	70	.015	7	Yes
28	2015	Berkhemer	127	500	.001	15	Yes
29	2016	Bracard	191	402	.036	2	Yes
30	2016	Li	126	210	<.001	12	Yes
31	2016	_koloud_k	95	242	.018	4	No
32	2017	Cannon	415	1962	<.001	51	Yes
33	2017	Oskouie	96	144	.002	7	No
34	2018	Nogueira	65	206	<.001	18	Yes
35	2018	Haesebaert	205	691	.016	6	Yes

Summary of the studies included in the analysis.[11-35] Overall, the median FI was 9 (IQR: 4-15.0), the mean FI was 11.36, and the studies with an FI less than or equal to 2 (lower FI) comprised 20% of the total (5 of 25).

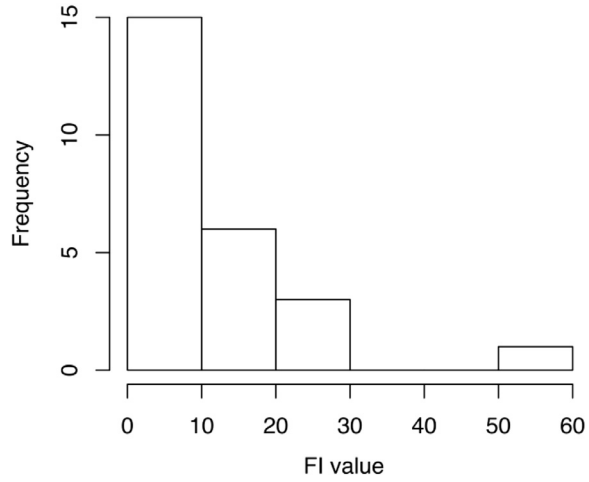

Figure 1. *Histogram of FI distribution in the current study (total n = 25). The FI distribution is plotted on the histogram. FI: median 7 (IQR: 4-15), mean 10.9.*

median total number of events was 73 (IQR: 25-127). There was no significant correlation between the FI and total number of participants (P = .059, Spearman's rank correlation), but there was a significant correlation between the FI and total number of events in each study (rho = .430, P = .032). Of all the studies, 48% (12 of 25) were associated with acute settings 48%, while the remaining 52% (13 of 25) were associated with nonacute settings. There was no significant difference in the FI (P = .172, Wilcoxon rank sum test) or total sample size (P = .644, Wilcoxon rank sum test) between RCTs with acute settings and those without.

In the review of earlier literature evaluating the FI in other clinical fields, 10 studies were identified (Table 2).

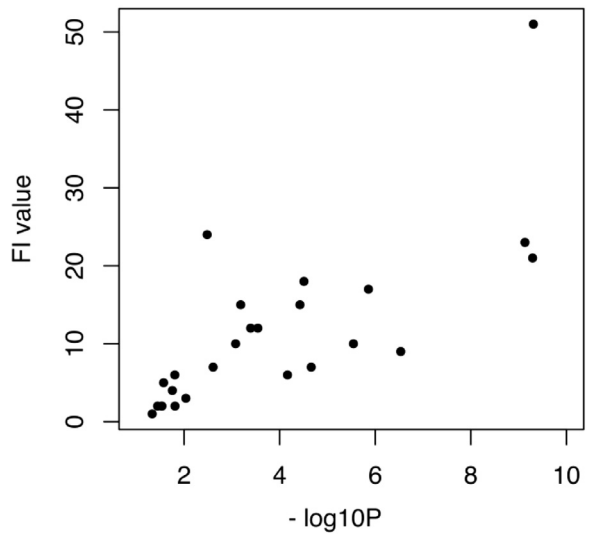

Figure 2. *Scatter graph of FI to the observed P value logarithm. The scatter graph of FI to the observed P value logarithm in the included studies shows a strong positive correlation (rho = .824, P < .001, Spearman's rank correlation).*

These included a median 47.5 studies (IQR: 30.3-92) and a median 141 (IQR: 100-161.8) total number of participants in each study. The median FI was 2.5 (IQR: 2-7.8), which is in contrast to our study, where only 20.0% (5 of 25) of studies included had an FI of 2.5 or less. The median FI and median number of included studies in each field combined with our study results are plotted in Figure 3. Although our analysis has a relatively smaller median number of included studies, the mean FI, weighted with the number of included studies, was 6.2 in past FI studies of other clinical fields, which is still lower than our study result (mean 10.9).

Discussion

In general, our present study replicated the results as of previous FI studies.[1-10] For example, the FI distribution in this study resembled that seen in earlier FI studies, which obeys a Poisson distribution. The observed P value and FI showed a strong negative correlation, as should be deducted by the calculation of FI. Finally, the FI was significantly associated with the total number of events, but not with the total sample size.

Our study reported that the median FI was 7, which indicates that at least 7 participants in 1 arm of the RCT are required to change from a nonevent to an event in order to change the statistically significant result to nonsignificant. In contrast, the median value of the FI in previous FI studies was 2.5.[1-10] In previous studies which evaluated the FI of other clinical fields, more than half (6 of 10) of the studies reported a median FI of less than or equal to 2, whereas the FI in our study is in the upper half of the FI distribution. The mean FI in our study was also higher than the weighted mean of the FI in previous studies. These results suggest that many RCTs in the field of stroke may have a fair robustness compared to those in other clinical fields.

Our study has some limitations. First, as mentioned in the Methods section, we have not evaluated the "loss to follow-up" factor, which can be useful in evaluating the FI in RCTs. For example, the result significance can be easily influenced if the FI is equal to or more than the "loss to follow-up" number. In addition, the mere comparison of representative values (ie, median FI) between our results and those of previous FI studies may be insufficient. Furthermore, since the literature review in this study is not always completely systematic, there may be a selection bias as to the clinical field of "ischemic stroke."

To conclude, we applied a new metric for ischemic stroke RCTs to evaluate their robustness, which suggested that many RCTs in this field have a fair robustness. Although the true significance and validity in using FI remains unclear, the FI can be an option as an easy-to-understand statistical metric for the results of RCTs with dichotomous endpoints.

Table 2. *Previous studies evaluating the FI in RCTs of other clinical fields*

Ref#	Year	Author	Field	Included studies	FI_median	FI_IQR	Total sample size (median)	Total sample size (IQR)
[1]	2014	Walsh	Any	399	8	(NA-NA)	682	(NA-NA)
[2]	2015	Evniew	Spine_surgery	40	2	(1-3)	132	(79-208)
[3]	2016	Ridgeon	Critical_care	56	2	(1-3.5)	126.5	(87-326)
[4]	2017	Matics	Pediatric	17	7	(2-11)	152	(NA-NA)
[5]	2017	Khan	Sports_surgery	48	2	(1-2.8)	64	(48.5-89.5)
[6]	2017	Shen	ICH	47	2	(1-4)	165	(87-200)
[7]	2017	Docherty	Cardiology	25	26	(NA-NA)	2331	(NA-NA)
[8]	2018	Shen	Ophthalmology	156	2	(NA-NA)	91.5	(NA-NA)
[9]	2018	Mazzinari	Anestheology	104	3	(2-7)	150	(70-300)
[10]	2018	Noel	Head_and_neck_surgery	27	8	(2.3-18.3)	67.5	(42-143)
		Current study	Ischemic stroke	25	9	(4-15)	206	(144-450)

Summary of 10 earlier studies evaluating the FI in other clinical fields.[1-10] They included median 43.5 (IQR: 29-54) of RCT studies with median 141 (IQR: 100-161.8) of total sample size. The median FI was 2.5 (IQR: 2-7.8), which is in contrast to our study, where only 20.0% (5 of 25) of studies included in ours had an FI of 2.5 or less.

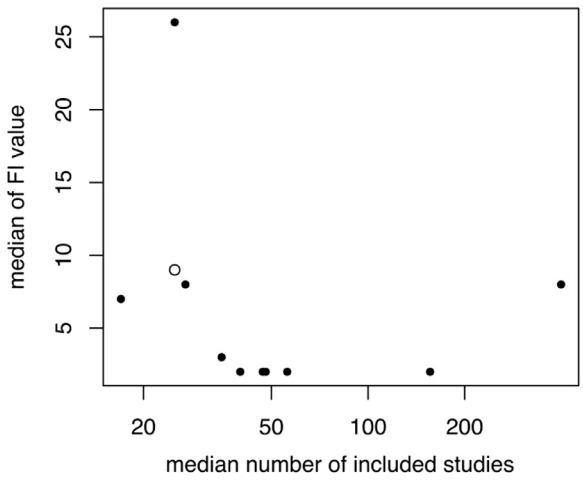

Figure 3. *Scatter graph of the median FI to the median number of included studies both in previous literature and the present study. Scatter graph of the median FI to the median number of included studies both previous literature and present study. Black circles denote results of previous literature, while a white circle denotes a result of the present study.*

Conflicts of Interest

The authors have no conflicts of interest to disclose.

References

1. 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.
2. 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.
3. Ridgeon EE, Young PJ, Bellomo R, et al. The Fragility Index in multicenter randomized controlled critical care trials. Crit Care Med 2016;44:1278-1284.
4. Matics TJ, Khan N, Jani P, et al. The Fragility Index in a cohort of pediatric randomized controlled trials. J Clin Med 2017;6. pii: E79.
5. 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.
6. Shen Y, Cheng X, Zhang W. The fragility of randomized controlled trials in intracranial hemorrhage. Neurosurg Rev 2017. https://doi.org/10.1007/s10143-017-0870-8. [Epub ahead of print].
7. Docherty KF, Campbell RT, Jhund PS, et al. How robust are clinical trials in heart failure? Eur Heart J 2017;38:338-345.
8. Shen C, Shamsudeen I, Farrokhyar F, et al. Fragility of results in ophthalmology randomized controlled trials: a systematic review. Ophthalmology 2018;125:642-648.
9. 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.
10. Noel CW, McMullen C, Yao C, et al. The fragility of statistically significant findings from randomized trials in head and neck surgery. Laryngoscope 2018. https://doi.org/10.1002/lary.27183. [Epub ahead of print].
11. Rosso C, Corvol JC, Pires C, et al. Intensive versus subcutaneous insulin in patients with hyperacute stroke: results from the randomized INSULINFARCT trial. Stroke 2012;43:2343-2349.
12. Asteriou C, Antonitsis P, Argiriadou H, et al. Minimal extracorporeal circulation reduces the incidence of postoperative major adverse events after elective coronary artery bypass grafting in high-risk patients. A single-institutional prospective randomized study. Perfusion. 2013;28:350-356.
13. Cao YJ, Zhang X, Wang WH, et al. Oral fibrinogen-depleting agent lumbrokinase for secondary ischemic stroke prevention: results from a multicenter, randomized, parallel-group and controlled clinical trial. Chin Med J (Engl) 2013;126:4060-4065.
14. Higgins P, MacFarlane PW, Dawson J, et al. Noninvasive cardiac event monitoring to detect atrial fibrillation after ischemic stroke: a randomized, controlled trial. Stroke 2013;44:2525-2531.
15. Reddy VY, Sievert H, Halperin J, et al. PROTECT AF steering committee and investigators. Percutaneous left atrial appendage closure vs warfarin for atrial fibrillation: a randomized clinical trial. JAMA 2014;312:1988-1998.

16. Zhang C, Zhang R, Zhang S, et al. Baclofen for stroke patients with persistent hiccups: a randomized, double-blind, placebo-controlled trial. Trials 2014;15:295.
17. Di Biase L, Burkhardt JD, Santangeli P, et al. Periprocedural stroke and bleeding complications in patients undergoing catheter ablation of atrial fibrillation with different anticoagulation management: results from the role of Coumadin in preventing thromboembolism in atrial fibrillation (AF) patients undergoing catheter ablation (COMPARE) randomized trial. Circulation 2014;129:2638-2644.
18. Piironen K, Tiainen M, Mustanoja S, et al. Mild hypothermia after intravenous thrombolysis in patients with acute stroke: a randomized controlled trial. Stroke 2014;45:486-491.
19. Suntrup S, Marian T, Schröder JB, et al. Electrical pharyngeal stimulation for dysphagia treatment in tracheotomized stroke patients: a randomized controlled trial. Intensive Care Med 2015;41:1629-1637.
20. Zaidat OO, Fitzsimmons BF, Woodward BK, et al. Effect of a balloon-expandable intracranial stent vs medical therapy on risk of stroke in patients with symptomatic intracranial stenosis: the VISSIT randomized clinical trial. JAMA 2015;313:1240-1248.
21. Kuliha M, Roubec M, Procházka V, et al. Randomized clinical trial comparing neurological outcomes after carotid endarterectomy or stenting. Br J Surg 2015;102:194-201.
22. Gong J, Chen X, Li S. Efficacy of a community-based physical activity program KM2H2 for stroke and heart attack prevention among senior hypertensive patients: a cluster randomized controlled phase-II trial. PLoS One 2015;10:e0139442.
23. Huo Y, Li J, Qin X, et al. Efficacy of folic acid therapy in primary prevention of stroke among adults with hypertension in China: the CSPPT randomized clinical trial. JAMA 2015;313:1325-1335.
24. Jovin TG, Chamorro A, Cobo E, et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. N Engl J Med 2015;372:2296-2306.
25. Saver JL, Goyal M, Bonafe A, et al. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. N Engl J Med 2015;372:2285-2295.
26. Goyal M, Demchuk AM, Menon BK, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. N Engl J Med 2015;372:1019-1030.
27. Campbell BC, Mitchell PJ, Kleinig TJ, et al. Endovascular therapy for ischemic stroke with perfusion-imaging selection. N Engl J Med 2015;372:1009-1018.
28. Berkhemer OA, Fransen PS, Beumer D, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. N Engl J Med 2015;372:11-20.
29. Bracard S, Ducrocq X, Mas JL, et al. Mechanical thrombectomy after intravenous alteplase versus alteplase alone after stroke (THRACE): a randomised controlled trial. Lancet Neurol 2016;15:1138-1147.
30. Li JY, Yuan LX, Zhang GM, et al. Activating blood circulation to remove stasis treatment of hypertensive intracerebral hemorrhage: a multi-center prospective randomized open-label blinded-endpoint trial. Chin J Integr Med 2016;22:328-334.
31. Školoudík D, Kuliha M, Hrbáč T, et al. Sonolysis in prevention of brain infarction during carotid endarterectomy and stenting (SONOBUSTER): a randomized, controlled trial. Eur Heart J 2016;37:3096-3102.
32. Cannon CP, Bhatt DL, Oldgren J, et al. Dual antithrombotic therapy with dabigatran after PCI in atrial fibrillation. N Engl J Med 2017;377:1513-1524.
33. Savadi Oskouie D, Sharifipour E, Sadeghi Bazargani H, et al. Efficacy of citalopram on acute ischemic stroke outcome: a randomized clinical trial. Neurorehabil Neural Repair 2017;31:638-647.
34. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. N Engl J Med 2018;378:11-21.
35. Haesebaert J, Nighoghossian N, Mercier C, et al. Improving access to thrombolysis and inhospital management times in ischemic stroke: a stepped-wedge randomized trial. Stroke 2018;49:405-411.