The Stata Journal (2022)
22, Number 1, pp. 77–88 DOI: 10.1177/1536867X221083856

Computing the fragility index for randomized trials and meta-analyses using Stata

Ariel Linden
Linden Consulting Group
San Francisco, CA
alinden@lindenconsulting.org

Abstract. In this article, I introduce two commands for computing the fragility index (FI): `fragility`, which is used for individual randomized controlled trials, and `metafrag`, which is used for meta-analyses. The FI for individual studies is defined as the minimum number of patients whose status would have to change from a nonevent to an event to nullify a statistically significant result. Correspondingly, the FI for meta-analyses is defined as the minimum number of patients from one or more trials included in the meta-analysis for which a modification of the event status (that is, changing events to nonevents or nonevents to events) would change the statistical significance of the pooled treatment effect to nonsignificant. Whether for an individual study or for a meta-analysis, a low FI indicates a more "fragile" study result, and a larger FI indicates a more robust result.

Keywords: st0664, fragility, metafrag, fragility index, meta-analysis, randomized controlled trials, research methodology, statistical significance

1 Introduction

When considering the results of a randomized controlled trial (RCT), scientists and those who rely on scientific evidence often conclude that a treatment is effective solely based on a p-value threshold (that is, < 0.05). However, the use of a p-value threshold to declare statistical significance has been widely criticized for being overly simplistic, frequently misunderstood, and inappropriately interpreted (see, for example, Amrhein, Greenland, and McShane [2019]; Colquhoun [2017]; Feinstein [1998]; Ioannidis [2005, 2018]; Sterne and Davey Smith [2001]; Wasserstein and Lazar [2016]).

As an upshot of this discourse, several supplementary measures to the p-value have been proposed to provide more focus on the robustness of statistically significant results from RCTs. Among these are Bayesian analyses (Quatto, Ripamonti, and Marasini 2020); the type S ("sign") error risk and exaggeration ratio (Gelman and Tuerlinckx 2000; Gelman and Carlin 2014); S-values (Greenland 2019); second-generation p-values (Blume et al. 2019); and the fragility index (FI) (Walsh et al. 2014).

In this article, I introduce two commands for computing the FI: the `fragility` command, which is used for individual RCTs with a binary outcome (Walsh et al. 2014), and the `metafrag` command for meta-analysis with a binary outcome (Atal et al. 2019). For single studies, the FI is defined as the minimum number of patients whose status would have to change from a nonevent to an event to nullify a statistically significant result. A smaller FI indicates that the statistical significance is contingent on only a small number of events, whereas a larger FI indicates a more robust result. The FI for meta-analysis is defined as the minimum number of patients from one or more trials included in the meta-analysis for which a modification of the event status (that is, changing events to nonevents or nonevents to events) would change the statistical significance of the pooled treatment effect to nonsignificant (Atal et al. 2019). As such, an FI of zero indicates that no modification of the event status is necessary to elicit a statistically nonsignificant pooled treatment effect. Conversely, a large FI score indicates that many modifications to the event status are required to change a statistically significant pooled effect to nonsignificant (and thus, the results may be considered more robust).

2 Methods

2.1 Computing the FI for individual RCTs

The FI represents the absolute number of additional events (primary endpoints) required to obtain a p-value greater than or equal to a predetermined statistical significance threshold (typically set to 0.05). The FI for individual RCTs is computed by adding an event to the study group with the smaller number of events (and subtracting a nonevent from the same group to keep the total number of patients within that group constant) and recomputing the two-sided significance. Events are iteratively added until the first time the computed p-value becomes statistically nonsignificant (Walsh et al. 2014).

`fragility` also computes the fragility quotient as proposed by Ahmed, Fowler, and McCredie (2016). The fragility quotient is a relative measure of fragility that simply divides the absolute FI by the total sample size (Ahmed, Fowler, and McCredie 2016).

2.2 Computing the fragility index for meta-analyses

To evaluate the FI of a meta-analysis, one sequentially recalculates the 95% confidence interval (CI) of the pooled estimate after performing all single event-status modifications that increase the estimate (or decrease it, depending on whether the treatment is expected to increase or decrease the risk of the outcome) by 1) changing a nonevent to an event for patients receiving treatment A for each single trial or 2) changing an event to a nonevent for patients receiving treatment B for each trial (Atal et al. 2019).

This process leads to $2N$ newly calculated 95% CIs for the pooled estimate (where N is the total number of studies in the meta-analysis). If one of the newly calculated CIs overlaps 1.0, the FI of the meta-analysis is 1 because one unique event-status modification (that is, changing a nonevent to an event in arm A or an event to a nonevent in arm B) in one specific trail changed the statistical significance of the meta-analysis. If all the newly calculated 95% CIs for the pooled estimate remain < 1.0 (in the case of a treatment that lowers the risk of the outcome or > 1.0 if the treatment is expected to increase the probability of the outcome), the specific trial and specific event-status modification that lead to the 95% CI for the pooled estimate being closer to 1.0 as a starting point for the next iteration are selected (Atal et al. 2019).

This process is then repeated by performing a new single event-status modification in each arm of each trial in turn on top of the first selected modification. Similarly, if one of these $2N$ event-status modifications leads to a newly calculated 95% CI for the pooled estimate overlapping 1.0, the FI of the meta-analysis is then equal to 2. This process is iterated until one event-status modification leads to a newly calculated 95% CI for the pooled estimate overlapping 1.0. The number of iterations needed to find a combination of event-status modifications in specific arms and trials leading to a modified meta-analysis with 95% CI for the pooled estimate overlapping 1.0 is thus the FI for the meta-analysis (Atal et al. 2019).

2.3 Differences between metafrag and the R package fragility_ma

`metafrag` produces results consistent with those of the R package `fragility_ma` and its related website http://www.clinicalepidemio.fr/fragility_ma/. However, there are some differences between the software programs: 1) Stata's `meta esize` command does not support the combination of random effects with the Mantel–Haenszel method (see `help meta_esize##remethod`), whereas `fragility_ma`, which uses the R package `metabin` for computing pooled treatment effects, does support this combination; 2) Stata's `meta esize` handles zero cells somewhat differently from `metabin`, possibly leading to slightly different results between software packages when some individual studies have zero cells; and 3) when there are ties between studies in the computed maximum (minimum) confidence level at any iteration, `fragility_ma` reports the FI that includes the modifications to all tied studies. `metafrag` reports both the FI for each iteration in the loop where any event modification occurs and the total number of modifications if there are ties.

3 The fragility command

This section describes the syntax of the `fragility` command and available options. `fragility` is an immediate command (see [U] **19 Immediate commands**).

3.1 Syntax

```
fragility #n11 #n12 #n21 #n22 [, level(#) chi2 detail]
```

In the syntax, variables *#n11* and *#n12* contain the respective numbers of events and nonevents from individuals in group 1 (treatment), and variables *#n21* and *#n22* contain the respective numbers for group 2 (control).

3.2 Options

`level(#)` specifies the desired p-value threshold level at which to test statistical significance. Most disciplines tend to use the p-value threshold of 0.05 to imply that the observed result is unlikely to occur by chance. However, some disciplines set the threshold for statistical significance more liberally to 0.10, while others may set the threshold more conservatively, such as to 0.01. `level(#)` allows users to set their own threshold. The default is `level(0.05)`.

`chi2` calculates and displays Pearson's χ^2 for the hypothesis that the rows and columns in a two-way table are independent. The default is Fisher's exact test, which generally produces more conservative estimates.

`detail` displays all the 2×2 tables produced during the iterative process of adding events to the group with the lowest actual number of events until the *p*-value threshold is met or surpassed.

3.3 Stored results

`fragility` stores the following in `r()`:

Scalars	
`r(fi)`	FI
`r(fq)`	fragility quotient
`r(pval)`	*p*-value at the FI

4 The metafrag command

This section describes the syntax of the `metafrag` command and available options. `metafrag` is a postestimation command for `meta esize` (see [META] **meta esize**), thereby capitalizing on the comprehensive list of options available in official Stata's `meta` suite for computing effect sizes for binary outcomes.

4.1 Syntax

`metafrag` [`,` `eform` `forest`[`(`*forestplot*`)`]]

4.2 Options

`eform` reports exponentiated effect sizes and transforms their respective CIs whenever applicable. By default, the results are displayed in the metric declared with `meta esize` such as log odds-ratios and log risk-ratios (RRs). `eform` uses odds ratios when used with log odds-ratios declared with `meta esize` or RRs when used with the declared log RRs. `eform` affects how results are displayed, not how they are estimated and stored.

`forest`[`(`*forestplot*`)`] displays a forest plot of the studies after modification to the events and nonevents of included studies to move the pooled effect from statistically significant to nonsignificant (the user can set the level that "significance" represents using the `level()` option in `meta esize`). Specifying `forest` without options uses the default forest plot settings (with only the column headers modified). Studies that have event modifications are highlighted in blue (when events are added) and red (when events are subtracted).

4.3 Stored results

`metafrag` stores the following in `r()`:

Scalars

`r(frag)`	FI for meta-analysis
`r(frag_ties)`	FI for meta-analysis when there are ties
`r(changes)`	number of studies in which events were modified

5 Examples

In this section, we demonstrate the use of `fragility` with two artificial examples and the use of `metafrag` with two empirical examples. For both commands, the first example illustrates the case of a fragile study result, and the second illustrates a more robust result. For the `metafrag` examples, the presented data correspond with real meta-analyses from Cochrane Systematic Reviews. The measures used for evaluating the treatment effect and for deriving the pooled treatment effects were the same as those used in the original Cochrane Systematic Reviews.

5.1 A fragile RCT

This example from Walsh et al. (2014) specifies that group 1 has 1 event and 99 nonevents and group 2 has 9 events and 91 nonevents.

```
. fragility 1 99 9 91

   Fragility index: 1
   Fragility quotient: 0.005
   p-value (exact): 0.058

   A fragility index of 1 indicates that group 1
   would require 1 additional events to obtain
   a p-value >= 0.050 using Fisher's exact test.
```

As shown in the output, the resulting FI of 1 suggests that the inference of a treatment effect is “fragile.” That is, only one additional event is needed to flip the results from being statistically significant to nonsignificant at the 0.05 level.

5.2 A more robust RCT

In example 2 from Walsh et al. (2014), group 1 has 200 events and 3,800 nonevents, and group 2 has 250 events and 3,750 nonevents:

```
. fragility 200 3800 250 3750

  Fragility index: 9
  Fragility quotient: 0.001
  p-value (exact): 0.054

  A fragility index of 9 indicates that group 1
  would require 9 additional events to obtain
  a p-value >= 0.050 using Fisher's exact test.
```

As shown in the output, the resulting FI of 9 suggests that the inference of a treatment effect is more robust than that of example 1.

5.3 A fragile meta-analysis

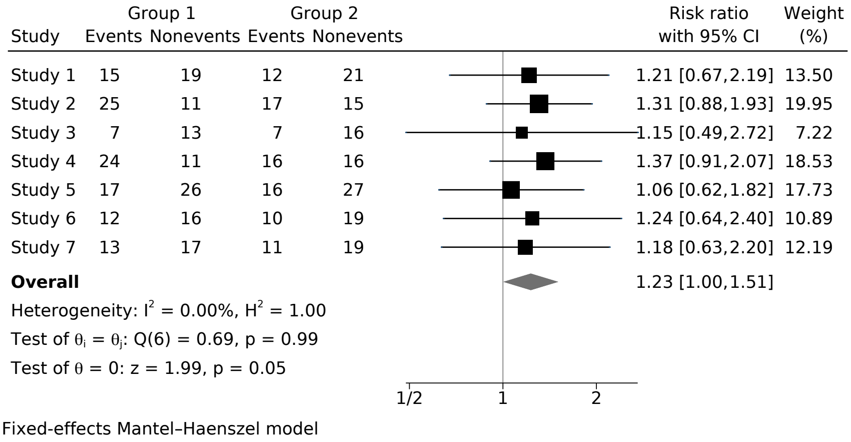
This meta-analysis includes 7 individual studies, with a total of 448 patients. We first load the data and then use `meta esize` to compute and declare effect sizes for a two-group comparison of binary outcomes. The log RR is specified as the effect size, and the fixed-effects meta-analysis is specified using the Mantel–Haenszel method.

```
. use example1
. meta esize events_1 noevents_1 events_2 noevents_2, esize(lnrratio) fixed(mh)
  (output omitted)
```

Next, we plot a forest plot of these data, specifying that the results be presented as exponentiated values, and modify some elements of the display (see [META] **meta forestplot**):

```
. meta forestplot, eform nullrefline columnopts(_data1, supertitle(Group 1))
> columnopts(_data2, supertitle(Group 2))
> columnopts(_a _c, title("Events"))
> columnopts(_b _d, title("Nonevents"))

  Effect-size label: Log risk-ratio
        Effect size: _meta_es
          Std. err.: _meta_se
```



Study	Group 1 Events	Group 1 Nonevents	Group 2 Events	Group 2 Nonevents	Risk ratio with 95% CI	Weight (%)
Study 1	15	19	12	21	1.21 [0.67,2.19]	13.50
Study 2	25	11	17	15	1.31 [0.88,1.93]	19.95
Study 3	7	13	7	16	1.15 [0.49,2.72]	7.22
Study 4	24	11	16	16	1.37 [0.91,2.07]	18.53
Study 5	17	26	16	27	1.06 [0.62,1.82]	17.73
Study 6	12	16	10	19	1.24 [0.64,2.40]	10.89
Study 7	13	17	11	19	1.18 [0.63,2.20]	12.19
Overall					1.23 [1.00,1.51]	

Heterogeneity: $I^2 = 0.00\%$, $H^2 = 1.00$

Test of $\theta_i = \theta_j$: Q(6) = 0.69, p = 0.99

Test of $\theta = 0$: z = 1.99, p = 0.05

Fixed-effects Mantel–Haenszel model

As shown in the forest plot, the treatment was associated with a statistically significant increase in the risk of the outcome (RR 1.23, 95% CI [1.00 to 1.51]). Next we use `metafrag` to compute the FI and specify the options `forest` and `eform`:

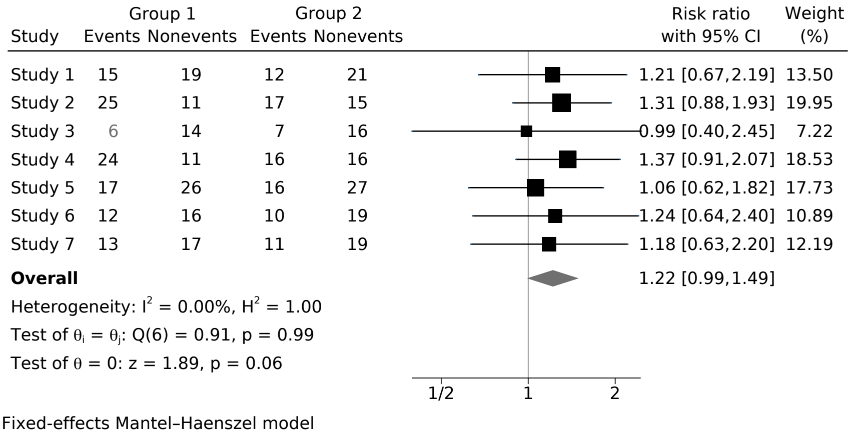
```
. metafrag, forest eform

Computing the fragility index. Please wait...
────┼──── 1 ────┼──── 2 ────┼──── 3 ────┼──── 4 ────┼──── 5

   Fragility Index: 1

   The pooled treatment effect turns statistically nonsignificant
   after 1 event-status modifications

   1 trial was modified:
   -  Study 3: subtracted 1 event from Group 1
```



Study	Group 1 Events	Group 1 Nonevents	Group 2 Events	Group 2 Nonevents	Risk ratio with 95% CI	Weight (%)
Study 1	15	19	12	21	1.21 [0.67,2.19]	13.50
Study 2	25	11	17	15	1.31 [0.88,1.93]	19.95
Study 3	6	14	7	16	0.99 [0.40,2.45]	7.22
Study 4	24	11	16	16	1.37 [0.91,2.07]	18.53
Study 5	17	26	16	27	1.06 [0.62,1.82]	17.73
Study 6	12	16	10	19	1.24 [0.64,2.40]	10.89
Study 7	13	17	11	19	1.18 [0.63,2.20]	12.19
Overall					1.22 [0.99,1.49]	

Heterogeneity: $I^2 = 0.00\%$, $H^2 = 1.00$

Test of $\theta_i = \theta_j$: Q(6) = 0.91, p = 0.99

Test of $\theta = 0$: z = 1.89, p = 0.06

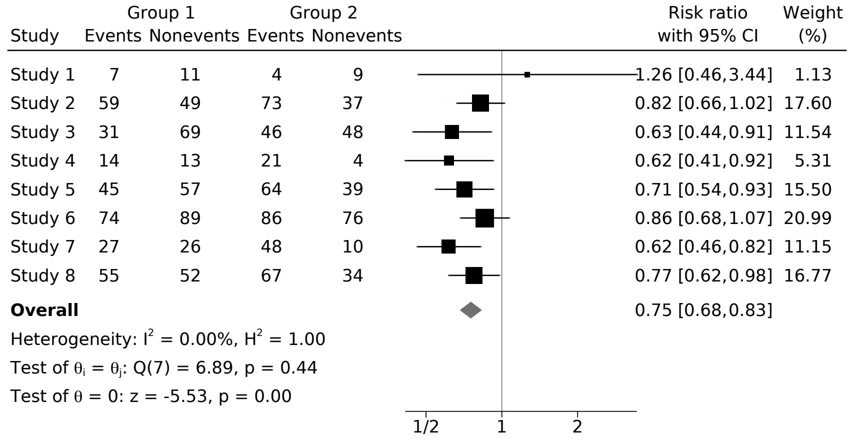
Fixed-effects Mantel–Haenszel model

As shown in the output, the FI is 1, indicating that the pooled treatment effect turns statistically nonsignificant after only one event-status modification. In this meta-analysis, the one event modification was made by subtracting one event from group 1 in study 3. In the forest plot, this addition corresponds with the value highlighted in gray (red on actual screen) under group 1 in study 3. The RR for the pooled effect is now statistically nonsignificant ([RR] 1.22, 95% CI [0.99 to 1.49]).

5.4 A more robust meta-analysis

This meta-analysis includes 8 individual studies, with a total of 1,344 patients. As before, we first load the data, then use `meta esize` to compute and declare effect sizes for a two-group comparison of binary outcomes, and then plot the forest plot:

```
. use example2, clear
. meta esize events_1 noevents_1 events_2 noevents_2, esize(lnrratio) fixed(mh)
  (output omitted)
. meta forestplot, eform nullrefline columnopts(_data1, supertitle(Group 1))
> columnopts(_data2, supertitle(Group 2))
> columnopts(_a _c, title("Events"))
> columnopts(_b _d, title("Nonevents"))
  Effect-size label: Log risk-ratio
        Effect size: _meta_es
          Std. err.: _meta_se
```



	Group 1		Group 2		Risk ratio	Weight
Study	Events	Nonevents	Events	Nonevents	with 95% CI	(%)
Study 1	7	11	4	9	1.26 [0.46,3.44]	1.13
Study 2	59	49	73	37	0.82 [0.66,1.02]	17.60
Study 3	31	69	46	48	0.63 [0.44,0.91]	11.54
Study 4	14	13	21	4	0.62 [0.41,0.92]	5.31
Study 5	45	57	64	39	0.71 [0.54,0.93]	15.50
Study 6	74	89	86	76	0.86 [0.68,1.07]	20.99
Study 7	27	26	48	10	0.62 [0.46,0.82]	11.15
Study 8	55	52	67	34	0.77 [0.62,0.98]	16.77
Overall					0.75 [0.68,0.83]	

Heterogeneity: I^2 = 0.00%, H^2 = 1.00

Test of $\theta_i = \theta_j$: Q(7) = 6.89, p = 0.44

Test of θ = 0: z = -5.53, p = 0.00

Fixed-effects Mantel–Haenszel model

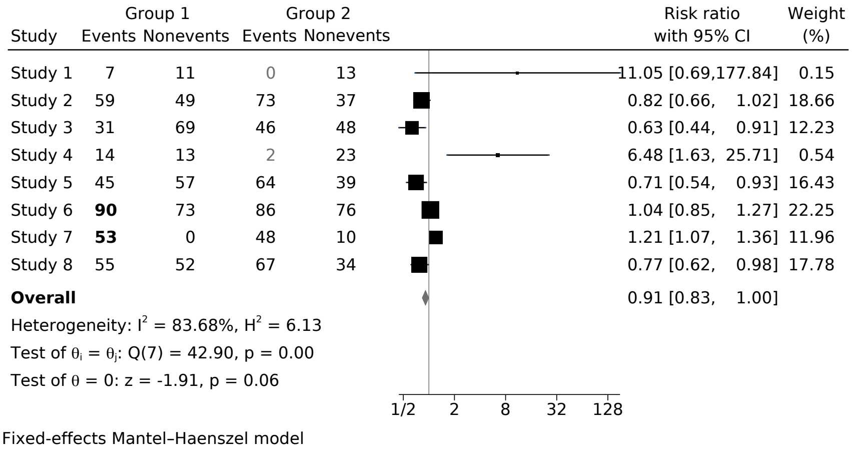
As shown in the forest plot, the treatment was associated with a statistically significant reduction in the risk of the outcome (RR 0.75, 95% CI [0.68 to 0.83]). Next, we use `metafrag` to compute the FI and specify the options `forest` and `eform`:

```
. metafrag, forest eform

Computing the fragility index. Please wait...
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
...............
   Fragility Index: 65

   The pooled treatment effect turns statistically nonsignificant
   after 65 event-status modifications

   A total of 4 trials were modified:
   -  Study 1: subtracted 4 events from Group 2
   -  Study 4: subtracted 19 events from Group 2
   -  Study 6: added 16 events to Group 1
   -  Study 7: added 26 events to Group 1
```



	Group 1		Group 2			
Study	Events	Nonevents	Events	Nonevents	Risk ratio with 95% CI	Weight (%)
Study 1	7	11	0	13	11.05 [0.69,177.84]	0.15
Study 2	59	49	73	37	0.82 [0.66, 1.02]	18.66
Study 3	31	69	46	48	0.63 [0.44, 0.91]	12.23
Study 4	14	13	2	23	6.48 [1.63, 25.71]	0.54
Study 5	45	57	64	39	0.71 [0.54, 0.93]	16.43
Study 6	**90**	73	86	76	1.04 [0.85, 1.27]	22.25
Study 7	**53**	0	48	10	1.21 [1.07, 1.36]	11.96
Study 8	55	52	67	34	0.77 [0.62, 0.98]	17.78
Overall					0.91 [0.83, 1.00]	

Heterogeneity: I^2 = 83.68%, H^2 = 6.13

Test of $\theta_i = \theta_j$: Q(7) = 42.90, p = 0.00

Test of θ = 0: z = -1.91, p = 0.06

Fixed-effects Mantel–Haenszel model

As shown in the output, the FI is 65, indicating that the pooled treatment effect turns statistically nonsignificant after 65 event-status modifications, with the event modifications occurring in 4 studies. In the forest plot, event additions correspond with values highlighted in bold (blue on actual screen), and event subtractions correspond with values highlighted in gray (red on actual screen). The RR is now statistically nonsignificant ([RR] 0.91, 95% CI [0.83 to 1.00]). The FI suggests that the pooled estimate from this meta-analysis is more robust than that in the previous example, where only one event modification was necessary to nullify the statistical significance of the pooled estimate.

6 Discussion

In this article, I introduced the `fragility` and `metafrag` commands, which compute the FI for individual randomized trials and meta-analyses with binary outcomes, respectively.

While the FI offers an intuitive supplemental measure to the *p*-value in interpreting the reliability of study findings, it has its critics. In particular, Carter, McKie, and Storlie (2017) illustrated a strong inverse relationship between the FI and the log10 of the *p*-value because both operate by decreasing the differences in response rates, resulting in a quantification of how extreme the observed trial results are relative to the null condition. Thus, as is true with *p*-values, the FI should not be misinterpreted as a measure of clinical effect. In other words, a higher FI should not be interpreted to imply greater clinical effect than a lower FI; rather, it simply illustrates the strength of the statistical significance itself (Brown et al. 2019; Narayan et al. 2018).

In conclusion, the `fragility` and `metafrag` commands provide a convenient method for evaluating the reliability of "statistical significance" in RCTs and meta-analyses. I advocate the reporting of the FI in conjunction with *p*-values and CIs to assist investigators and others in weighing the evidence for study robustness.

7 Acknowledgments

I thank John Moran for advocating that I write both of these commands and for providing me with many of the references used in the introduction. I also thank Ignacio Atal for his support in testing the results reported by `metafrag` to assess their consistency with his R package `fragility_ma` and website http://www.clinicalepidemio.fr/fragility_ma/, and I thank Houssein Assaad at StataCorp for providing details of how Stata and R differ in their respective computations for meta-analyses. Finally, I thank the chief editor and anonymous reviewer for providing helpful comments to improve the article and commands.

8 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 22-1
. net install st0664      (to install program files, if available)
. net get st0664          (to install ancillary files, if available)
```

9 References

Ahmed, W., R. A. Fowler, and V. A. McCredie. 2016. Does sample size matter when interpreting the fragility index? *Critical Care Medicine* 44: e1142–e1143. https://doi.org/10.1097/CCM.0000000000001976.

Amrhein, V., S. Greenland, and B. McShane. 2019. Scientists rise up against statistical significance. *Nature* 567: 305–307. https://doi.org/10.1038/d41586-019-00857-9.

Atal, I., R. Porcher, I. Boutron, and P. I. Ravaud. 2019. The statistical significance of meta-analyses is frequently fragile: Definition of a fragility index for meta-analyses.

Journal of Clinical Epidemiology 111: 32–40. https://doi.org/10.1016/j.jclinepi.2019.03.012.

Blume, J. D., R. A. Greevy, V. F. Welty, J. R. Smith, and W. D. Dupont. 2019. An introduction to second-generation p-values. *American Statistician* 73: 157–167. https://doi.org/10.1080/00031305.2018.1537893.

Brown, J., A. Lane, C. Cooper, and M. Vassar. 2019. The results of randomized controlled trials in emergency medicine are frequently fragile. *Annals of Emergency Medicine* 73: 565–576. https://doi.org/10.1016/j.annemergmed.2018.10.037.

Carter, R. E., P. M. McKie, and C. B. Storlie. 2017. The fragility index: A p-value in sheep's clothing? *European Heart Journal* 38: 346–348. https://doi.org/10.1093/eurheartj/ehw495.

Colquhoun, D. 2017. The reproducibility of research and the misinterpretation of p-values. *Royal Society Open Science* 4: 171085. https://doi.org/10.1098/rsos.171085.

Feinstein, A. R. 1998. P-values and confidence intervals: Two sides of the same unsatisfactory coin. *Journal of Clinical Epidemiology* 51: 355–360. https://doi.org/10.1016/s0895-4356(97)00295-3.

Gelman, A., and J. Carlin. 2014. Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science* 9: 641–651. https://doi.org/10.1177/1745691614551642.

Gelman, A., and F. Tuerlinckx. 2000. Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics* 15: 373–390. https://doi.org/10.1007/s001800000040.

Greenland, S. 2019. Valid p-values behave exactly as they should: Some misleading criticisms of p-values and their resolution with s-values. *American Statistician* 73: 106–114. https://doi.org/10.1080/00031305.2018.1529625.

Ioannidis, J. P. A. 2005. Why most published research findings are false. *PLOS Medicine* 2: e124. https://doi.org/10.1371/journal.pmed.0020124.

———. 2018. The proposal to lower p value thresholds to .005. *Journal of the American Medical Association* 319: 1429–1430. https://doi.org/10.1001/jama.2018.1536.

Narayan, V. M., S. Gandhi, K. Chrouser, N. Evaniew, and P. Dahm. 2018. The fragility of statistically significant findings from randomised controlled trials in the urological literature. *BJU International* 122: 160–166. https://doi.org/10.1111/bju.14210.

Quatto, P., E. Ripamonti, and D. Marasini. 2020. Best uses of p-values and complementary measures in medical research: Recent developments in the frequentist and Bayesian frameworks. *Journal of Biopharmaceutical Statistics* 30: 121–142. https://doi.org/10.1080/10543406.2019.1632874.

Sterne, J. A. C., and G. Davey Smith. 2001. Sifting the evidence—What's wrong with significance tests? *British Medical Journal* 322: 226–231. https://doi.org/10.1136/bmj.322.7280.226.

Walsh, M., S. K. Srinathan, D. F. McAuley, M. Mrkobrada, O. Levine, C. Ribic, A. O. Molnar, N. D. Dattani, A. Burke, G. Guyatt, L. Thabane, S. D. Walter, J. Pogue, and P. J. Devereaux. 2014. The statistical significance of randomized controlled trial results is frequently fragile: A case for a fragility index. *Journal of Clinical Epidemiology* 67: 622–628. https://doi.org/10.1016/j.jclinepi.2013.10.019.

Wasserstein, R. L., and N. A. Lazar. 2016. The ASA statement on *p*-values: Context, process, and purpose. *American Statistician* 70: 129–133. https://doi.org/10.1080/00031305.2016.1154108.

About the author

Ariel Linden is a health services researcher specializing in the evaluation of healthcare interventions and policy changes. He is both an independent consultant and a research scientist in the Department of Medicine at the University of California, San Francisco. Thus far, he has written 50 community-contributed packages for Stata.