Analyzing DECREASE trials for extent of data fabrication

Chris HJ Hartgerink, Gerben ter Riet, Marleen Kemper

09 March, 2016

The effectiveness of beta-blockers in preventing perioperative mortatility in non-cardiac surgery has been subject of discussion due to findings of misconduct.[1–4]. Perioperative mortality is the deathrate of patients during the period of the surgical procedure, typically including admission, anaesthesia, surgery, and recovery. The trials that were found to be harmed were the DECREASE trials[1–3]. Upon being discovered as invalid, the DECREASE studies were excluded from a 2014 meta-analysis[4]. The corrected analysis found that conclusions on the effectiveness of beta-blockers in perioperative mortality were reversed --- beta-blockers seemed to actually *increase* perioperative mortality instead of decrease it.

The committee reports about the DECREASE studies indicate that data fabrication was likely[1–3], but the extent of the data fabrication was not clearly indicated or deemed estimable. In this paper, we aim to estimate the extent of data fabrication in the DECREASE studies[5,6]. We expect to confirm the conclusions of the reports that data fabrication is likely.

To investigate the evidence for data fabrication in the DECREASE studies[5,6], we take three steps. First, we try to replicate the findings from the 2014 meta-analysis[4], which should be possible with the trial-event data provided in that paper. Second, we evaluate the veracity of the DECREASE studies (i.e., the probability that these study findings occurred assuming no data fabrication occurred). Third, we reverse this assumption and assume that data fabrication did occur and estimate how many data points would have to be fabricated to reproduce the results of the DECREASE studies. We expand on the methods for these three steps below.

Step 1: replicating Bouri et al. (2014)

We first replicate parts of Bouri et al. This helps ensure that (1) their results are reproducible and (2) we are using the correct estimates in the following steps. In order to replicate the results of the meta-analysis, we extract the trial-event data available from figures 2 and 3 in the original paper. More specifically, we extract the raw trial-event data for the 2 (control vs experimental) by 2 (event or no event) design, which we use to recompute the natural logarithm of the risk ratio and its standard error.

Using the notation of Table 1, the log of the risk ratio per study is computed as

$$ln(RR) = ln\left(\frac{d/(b+d)}{c/(a+c)}\right)$$

The standard error of this statistic is computed as

$$se_{ln(RR)} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

In case there is a zero-count in either d or c from Table 1, a half (i.e., .5) is added to each cell to ensure that computations are possible (i.e., a, b, c, and d). This is automatically done by Review Manager, the software used by Bouri et al., and is the default way of dealing with zero-counts in analyzing contingency tables[7].

Table 1. An example of clinical trial summary data used throughout this research.

	Control	Experimental
No event	a	b
Event	c	d

We use the R package metafor[8] to meta-analyze the risk ratios. First, we analyze all the non-DECREASE studies to verify the estimates in Figure 2 of Bouri and analyze all DECREASE studies to verify the estimates in Figure 3 of Bouri. Second, we meta-analyze all studies combined, including a dummy-predictor for the DECREASE and non-DECREASE studies to replicate results presented in Figure 4 of Bouri. These estimates are made with random-effects model using the REML estimation method. REML estimation has been indicated to be the most optimal estimation method for heterogeneity in a random effects model[9]. Prior to comparison of the estimates, results need to be exponentiated because the meta-analysis is conducted on log transformed risk ratios.

Checklist

1. Create datafile with the trial-event data from Bouri et al. (2014)
2. Code calculation of ln(RR) and standard error calculation in R
3. Implement the use of metafor with the rma() function, using random effects estimation method REML
4. Run DECREASE studies
5. Run non-DECREASE studies
6. Run meta-regression with dummy predictor comparing DECREASE and non-DECREASE studies

Step 2: evaluating the veracity of DECREASE studies

We estimate the probability of the results in the DECREASE studies based on the estimates from the non-DECREASE studies. In order to do this, we assume that the

non-DECREASE studies provide the best known representation of the true effect of beta-blockers on perioperative mortality. This is also known as estimating the veracity[10], which indicates the plausibility of the data in a random sampling context. We assume to estimate the true population distribution of effects, not perturbed by publication bias due to statistical significance, because a substantial number of nonsignificant effects are included in the dataset.

Based on the estimated mean effect and credible interval in the non-DECREASE studies, we compute the probability of the DECREASE trials. The estimates of the non-DECREASE studies are obtained from step 1, which include the estimated risk ratio, and the 95% credibility interval as provided by the package metafor. Based on this, we assume a normal distribution of population effects with the estimated effect as the mean of the distribution. The 95% credibility interval denotes the bounds of the normal distribution that covers 95% of the density. The standard deviation is calculated as the distance from the mean to either bound, divided by 1.96.

The two-sided probability of the DECREASE trials' findings serves as the probability that the DECREASE trials' findings occurred naturally. In other words, this is the estimated probability of the DECREASE trials without data fabrication, assuming that the non-DECREASE trials are genuine. The combined probability of the DECREASE trials is also calculated, by taking the product of the probability of the two independent studies.

Checklist

1. Note the mean estimated ln(RR), credible interval upper- and lower bound from step 1
2. Calculate standard deviation of the ln(RR)
3. Compute two-sided probability for both DECREASE studies individually (for ln(RR))
4. Compute joint probability of DECREASE studies by taking the product of the two-sided probabilities for the individual studies.

Step 3: estimating the amount of fabricated data

We estimate the number of data points fabricated in each DECREASE study individually (I and IV), if we were to assume data fabrication occurred. This step is carried out regardless of the results of step 2. This provides an indication of the severity of the data fabrication as concluded by the misconduct reports[1–4].

We first estimate the mortality probability in each condition (beta-blocker and control), for both the DECREASE- and non-DECREASE studies separately. To this end, we run four meta-analyses to estimate the ln(odds) of mortality per stratum. We apply the same methods as in step 1 to meta-analyze the individual effects (i.e., metafor package, REML random effects estimate, credible intervals). This results in eight meta-analyses to estimate the log odds. These separate analyses are necessary,

because step 1 only provides an estimate of the overall risk ratio, but not of the odds that an event occurs conditional on exposure (beta-blocker or control) and study set (DECREASE- versus non-DECREASE).

The log odds per study are calculated for each condition. For example, for the control condition, based on Table 1, it is calculated as

$$ln(odds)=ln\left(\frac{c/(a+c)}{1-(c/(a+c))}\right)$$

accompanied by the standard error

$$se_{ln(odds)}=\frac{1}{(a+c)(c/(a+c))(1-(c/(a+c)))}$$

For the beta-blocker condition c is replaced by d, and a is replaced by b. If there are no events in that condition (i.e., zero-count), a half is added to each cell (see also Step 1).[7] Subsequently, meta-analyses are conducted per condition (beta-blocker and control), for both the DECREASE- and non-DECREASE studies separately.

We apply the inversion method to estimate the number of data points fabricated in the DECREASE studies. To this end, we assume that the DECREASE data are in fact fabricated, opposed to Step 2 where we assumed they were genuine. The inversion method iteratively simulates data where X data points in that study are fabricated, where X varies from 0 through N. Each simulated sample is based on the effect distributions estimated earlier. This procedure is carried out for both DECREASE I and DECREASE IV. The analysis script will be made publicly available, and we include the step-by-step process here.

1. Take the entire sample from one DECREASE study, with the prespecified group sizes from that DECREASE study (e.g., 49 controls and 51 patients on beta-blockers, N=100).
2. Randomly assign X={0, 1, 2, ... , N} of the data points to be fabricated.
3. Draw an ln(odds) from the corresponding estimated effect distribution (i.e., fabricated/genuine control/beta-blocker condition).
4. Sample an event with the probability calculated from the ln(odds) from the previous step and note the data point 0 (no event) or 1 (event).
5. When all data points have been generated, estimate the ln(RR) in this simulated sample with metafor and REML estimation.
6. Note whether this ln(RR) is more likely under the non-DECREASE effect distribution (coded 0) or under the DECREASE effect distribution (coded 1), based on two-sided p-values (larger p-value is more probable) from the estimated effect distributions
7. Repeat k times (k = number of iterations; these have yet to be determined but will be at least 1,000)

8. Calculate proportions of iterations where estimate was more likely under fabrication than under genuine data.
9. Note for how many data points from the fabricated (DECREASE) distribution the proportion is closest to .025 and .975 to provide a 95% confidence interval of data points fabricated.

Checklist

1. Calculate the ln(odds) per condition for both non-DECREASE and DECREASE studies.
2. Calculate the standard error of the ln(odds) per effect.
3. Compute the four meta-analyses to estimate the ln(odds) and credible intervals for each of the 2 (beta-blocker; control) by 2 (non-DECREASE; DECREASE) design.
4. Create function to apply inversion method.
5. Apply inversion method to both DECREASE studies separately.

References

1. Onderzoekscommissie Wetenschappelijke Integriteit. *Onderzoek Naar Mogelijke Schending van de Wetenschappelijke Integriteit: Beknopte Versie*. Erasmus MC; 2011. https://web.archive.org/web/20151113121125/http://www.erasmusmc.nl/cs-research/bijlagen/integriteit/rapport-poldermans-2011.

2. Commissie Vervolgonderzoek 2012. *Rapport Vervolgonderzoek Naar Mogelijke Schending van de Wetenschappelijke Integriteit*. Erasmus MC; 2012. https://web.archive.org/web/20151027084205/http://www.erasmusmc.nl/5663/135857/3675250/3706798/erasmusmc.commissie.verv.onderzoek.2012.

3. Commissie Vervolgonderzoek Wetenschappelijke Integriteit 2013. *Rapport*. Erasmus MC; 2014.

4. Bouri S, Shun-Shin MJ, Cole GD, Mayet J, Francis DP. Meta-analysis of secure randomised controlled trials of β-blockade to prevent perioperative death in non-cardiac surgery. *Heart*. 2014;100(6):456-464. doi:http://doi.org/10.1136/heartjnl-2013-304262.

5. Dunkelgrun M, Boersma E, Schouten O, et al. Bisoprolol and fluvastatin for the reduction of perioperative cardiac mortality and myocardial infarction in intermediate-risk patients undergoing noncardiovascular surgery: A randomized controlled trial (DECREASE-IV). *Annals of surgery*. 2009;249(6):921-926. doi:http://doi.org/10.1097/SLA.0b013e3181a77d00.

6. Poldermans D, Boersma E, Bax JJ, et al. The effect of bisoprolol on perioperative mortality and myocardial infarction in High-Risk patients undergoing vascular surgery. *The New England journal of medicine*. 1999;341(24):1789-1794. doi:http://doi.org/10.1056/NEJM199912093412402.

7. Agresti A. *Categorical Data Analysis*. Hoboken, NJ: John Wiley & Sons Inc.; 2002.

8. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. 2010;36:1-48. http://www.jstatsoft.org/v36/i03/.

9. Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the Random-Effects model. *Journal of educational and behavioral statistics: a quarterly publication sponsored by the American Educational Research Association and the American Statistical Association*. 2005;30(3):261-293. doi:http://doi.org/10.3102/10769986030003261.

10. Peters CFW, Klaassen CAJ, Wiel MA van de. *Evaluating the Scientific Veracity of Publications by Dr. Jens Forster*. University of Amsterdam; 2015.