Randomized trials are frequently fragmented in multiple secondary publications

Shanil Ebrahima,b,c,d,e, Luis Montoya f, Mostafa Kamal el Din c, Zahra N. Sohanic, g, h, Arnav Agarwalth, Sheena Bancei, Juliann Saquibj, Nazmus Saquibk, John P.A. Ioannidisa,b,l,m,*

aStanford Prevention Research Center, Department of Medicine, Stanford University, 1265 Welch Road, Stanford, CA 94305, USA
bMeta-Research Innovation Center at Stanford (METRICS), School of Medicine, Stanford University, 1265 Welch Road, Stanford, CA 94305, USA
cDepartment of Clinical Epidemiology & Biostatistics, McMaster University, 1200 Main Street West, Hamilton, Ontario L8S 4K1, Canada
dDepartment of Anesthesia, McMaster University, 1200 Main Street West, Hamilton, Ontario L8S 4K1, Canada
ePrevention Lab and Systematic Overviews through advancing Research Technology (SORT), SickKids Research Institute, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada
fDepartment of Trauma and Orthopaedics, University Health Network, 399 Bathurst Street, Toronto, Ontario M5T 2S8, Canada
gPopulation Genomics Program, Faculty of Health Sciences, McMaster University, 1200 Main Street West, Hamilton, Ontario L8S 4K1, Canada
hFaculty of Medicine, 1 King’s College Circle, University of Toronto, Toronto, Ontario M5S 1A8, Canada
iDepartment of Applied Psychology and Human Development, Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, Ontario M5S 1V6, Canada
jDepartment of Family and Community Medicine, Qassim University, Buraydah, Saudi Arabia
kCollege of Medicine, Saluaiman AlRajhi Colleges, Al Bukayriyah, Saudi Arabia
lDepartment of Health Research and Policy, Stanford University School of Medicine, 150 Governor’s Ln, Stanford, CA 94305, USA
mDepartment of Statistics, Stanford University School of Humanities and Sciences, Building 1, Main Quad, Stanford, CA 94305, USA

Accepted 6 May 2016; Published online xxxx

Abstract

Objective: To assess the frequency and features of secondary publications of randomized controlled trials (RCTs).

Study Design and Setting: For 191 RCTs published in high-impact journals in 2009, we searched for secondary publications coauthored by at least one same author of the primary trial publication. We evaluated the probability of having secondary publications, characteristics of the primary trial publication that predict having secondary publications, types of secondary analyses conducted, and statistical significance of those analyses.

Results: Of 191 primary trials, 88 (46%) had a total of 475 secondary publications by 2/2014. Eight trials had >10 (up to 51) secondary publications each. In multivariable modeling, the risk of having subsequent secondary publications increased 1.32-fold (95% CI 1.05–1.68) per 10-fold increase in sample size, and 1.71-fold (95% CI 1.19–2.45) in the presence of a design article. In a sample of 197 secondary publications examined in depth, 193 tested different hypotheses than the primary publication. Of the 193, 43 tested differences between subgroups, 85 assessed predictive factors associated with an outcome of interest, 118 evaluated different outcomes than the original article, 71 had differences in eligibility criteria, and 21 assessed different durations of follow-up; 176 (91%) presented at least one analysis with statistically significant results.

Conclusions: Approximately half of randomized trials in high-impact journals have secondary publications published with a few trials followed by numerous secondary publications. Almost all of these publications report some statistically significant results.

Accepted 6 May 2016; Published online xxxx

Keywords: Secondary publications; Randomized controlled trial; Clinical trial; Individual patient data; Multiplicity; Secondary findings

http://dx.doi.org/10.1016/j.jclinepi.2016.05.016
0895-4356/© 2016 Published by Elsevier Inc.
1. Background

Randomized controlled trials (RCTs) are the gold standard for determining the effectiveness of treatments [1]. RCTs are a challenging endeavor given the rigor in designing and conducting trials, resources required [2], and the time taken to complete them [3]. Many researchers may find leading RCTs unattractive for their career, as RCTs typically result in a single published report and they are called anecdotally the deathbed of the assistant professor. However, with the pressure of publishing in academia, the paradigm of only one publication stemming from an RCT may be changing with authors performing secondary analyses on their trial data to publish additional articles.

Data sets from RCTs can be rich sources for secondary analyses [4]. Some organizations, such as the National Institutes of Health, have policies that encourage data sharing, specifically to support secondary analyses [5]. The vast majority of secondary analyses are however performed by the original authors of each trial [6,7]. Even when authors share data with other investigators, some original authors remain as coauthors in the resulting publications [4]. Secondary analyses can offer additional insights beyond the primary publication of the trial results. They can also provide to scientists, physicians, and the public more complete information about interventions [8]. However, there are also criticisms of secondary analyses: they may lack statistical power for new hypotheses [9], and multiple post hoc analyses (e.g., subgroups) can generate spurious misleading findings [8]. There are also issues with fragmenting results with “salami publications” [10,11], and even duplicate publication of RCTs has been described [12,13]. Fragmentation of the evidence across multiple articles may confuse readers, clinicians, and systematic reviewers.

To our knowledge, there has been no empirical evaluation of the frequency and features of secondary publications of individual RCTs. Here, we aim to assess the phenomenon and its implications at large scale, using a sample of RCTs published in high-impact journals. Specifically, we aim to assess how many secondary publications of RCTs are published by the authors of the primary publication; how soon they appear; what types of trials lead to more secondary publications; what are the reasons for completing secondary publications; and whether these secondary analyses claim statistically significant research findings.

2. Methods

2.1. Eligibility criteria

Using a previously constructed database of a random sample of 200 RCTs reporting on a primary outcome in high-impact journals in 2009 [14], eligible studies for our current evaluation include all published secondary publications coauthored by at least one of the same author(s) of the original primary trial publication and using individual-level data from the original trial.

2.2. Primary trials

Primary trials have been previously identified and used in a project assessing the prevalence and impact of adjustments in results of RCTs [14]. In brief, searches were made in PubMed for study type = randomized controlled trial for the 25 biomedical journals with highest impact factor (Journal Citation Reports 2009) that are also likely to publish RCTs: BMJ, American Journal of Psychiatry, American Journal of Respiratory Critical Care Medicine, Annals of Internal Medicine, Annals of Neurology, Archives of General Psychiatry, Archives of Internal Medicine, Blood, Brain, Circulation, European Heart Journal, Gastroenterology, Gut, Hepatology, Journal of Allergy and Clinical Immunology, Journal of the American College of Cardiology, Journal of Clinical Oncology, Journal of the National Cancer Institute, JAMA, Lancet, Lancet Infectious Diseases, Lancet Neurology, Lancet Oncology, New England Journal of Medicine, and PLoS Medicine. We included studies involving human participants and published in 2009. As previously stated [14], these 200 trials were randomly chosen with random numbers from a total of 684 articles of trials published in these journals in 2009.

In our previous evaluation [14], we had already excluded three articles that did not analyze primary outcomes between the study arms. For the current project, we also excluded articles that did present some analyses of primary trial outcomes but had already been preceded by an earlier publication of other primary outcomes; conversely, we did not exclude trials with preceding publications that did not present any primary analyses (e.g., design articles, baseline data reports, early results, other secondary analyses). We thus checked the cited references of each article and identified citations to previous relevant publications of the same trial. This process eliminated another eight articles. Two articles each reported the results from two separate trials and were considered separately. Thus, 191 trials were finally selected for evaluation (Appendix A at www.jclinepi.com).

2.3. Selection of secondary publications

We used the Thomson Reuters ISI Web of Science database to identify secondary publications for the 191 original primary trials. It would be extremely unlikely for a secondary publication not to cite the primary trial. We identified and recorded the total citations of each primary trial publication until February 2014. We refined the “times cited” results to include only articles that include as an author any of the original trial authors. We recorded the number of citations by articles that use individual-level data from the original trial and included as an author any of the authors of the original trial. For articles considered to potentially reflect secondary publications, the full text was examined to confirm eligibility.

To identify secondary publications published before the primary publication of the primary outcome, we reviewed
Moreover, six extractors, assessing each article in duplicate, recorded whether the secondary publication articles tested the same hypothesis (same PICO-T = patient population, intervention, comparator, outcome, and time point). If the same hypothesis was tested, the same extractors independently captured details of differences in statistical or other analytical methods, handling of missing data, intention-to-treat or on-treatment principle, and other aspects of the analysis. The same extractors independently also assessed whether the secondary analysis modified the interpretation/conclusions of the original article about which patients should be treated with the experimental intervention [15].

If different hypotheses were tested compared to the original article, extractors assessed whether these reflected subgroup analyses, predictive analyses, of differences in outcomes analyzed, eligibility criteria for the analysis, duration of follow-up, or any other aspect. Extractors independently assessed whether the secondary analyses in articles testing a different hypothesis showed statistically significant results (P < 0.05 or 95% confidence interval excluding the null).

### 2.5. Protocols

For the 191 primary trials, we had previously requested the protocols of the trials from study authors and 59 protocols were retrieved [14]. Therefore, we evaluated whether any secondary analyses were prespecified in these protocols.

### 2.6. Statistical analysis

We described data as proportions, means or medians, and range and interquartile range as appropriate.

Using Kaplan—Meier plots, we estimated the probability of having the first secondary publication as a function of time after the publication of the primary trial results. Preprimary secondary publications were not considered as events. We used proportional hazards models to assess whether characteristics of the primary trial publications affected the probability of secondary publication. One set of models considered only the first secondary publication, whereas another set of models captured recurrent events (i.e., all secondary publications) [16]. The following characteristics were tested: number of primary outcomes (one vs. more); multisite vs. single site; sample size (log-scaled); journal (general medical vs. specialty); industry funding (yes/no); whether there was a secondary publication published before the primary publication; presence of design article; and significance of at least one primary outcome. General medical journal included the following: The New England Journal of Medicine, The Lancet, Journal of American Medical Association, British Medical Journal, PLoS Medicine, and Canadian Medical Association Journal. We subsequently developed stepwise backward elimination multivariable models to evaluate the adjusted association
between the characteristics that were found to be nominally significant ($P < 0.05$) in our respective univariable analyses.

We calculated the proportion of secondary publications including each type of secondary analysis and the proportion with any statistically significant results in each category and overall. We also evaluated whether the secondary findings were concordant with the primary trial findings (in terms of presence or not of statistical significance) whenever the secondary article had the same PICO as the primary article but differed in duration of follow-up.

We completed analyses using SPSS Statistics software version 21.0.0 (IBM Corp., Armonk, NY) and R version 3.1.2 (R Development Core Team, Vienna, Austria). We did not require approval from a research ethics board as we were using published results.

3. Results

3.1. Search results

We identified and screened 21,729 citations to the 191 primary trial reports (median citations per trial: 57 [range: 2–1,111; interquartile range [IQR] 30–123]) for eligibility...
(Fig. 1). A total of 558 articles underwent full-text screening. Of those, we excluded 75 articles as ineligible and were unable to locate another 8. Our final sample was comprised of 475 secondary publications (Appendix B at www.jclinepi.com). As shown (Fig. 1), 103 (54%) of the primary articles had no secondary publication recorded, whereas 88 (46%) had at least one secondary publication. There were a few primary articles that had many secondary publications, but most had only 1 (n = 27 primary articles), 2 (n = 24 primary articles), or 3 (n = 7 primary publications). Appendix B at www.jclinepi.com provides the 197 secondary publications analyzed in further depth.

### 3.2. Study characteristics

One hundred thirty (68%) primary articles were multisite trials and 47 (25%) primary trials had >1 primary outcome. Twenty-three (12%) had design articles. More than 40% of the primary trials were published in general medical journals, most articles were published by authors from United States and continental Europe and approximately 40% were industry funded (Table 1).

Eighteen (9%) secondary publications were published before the primary publication, with postprimary secondary publications being evenly spread through the subsequent years. Only 5% of all secondary publications were published in general medical journals. Most secondary publications also came from authors from United States and continental Europe. Approximately 40% of the secondary publications were industry funded (Table 1).

Eight primary trials [17–24] had >10 secondary publications identified, with one trial [20] having 51 secondary publications. Four of the eight evaluated a population with cardiac disease, all were multisite trials, and most were funded by industry and had design articles (Table 2). These eight trials accounted for 54% (257/475) of the secondary publications.

Among the 197 secondary publications assessed in depth, 1 (0.5%) study tested the same hypothesis [25], 2 (1%) presented only baseline data [26,27], 1 (0.5%) presented the process of a quality assurance program [28], and 193 (98%) tested different hypotheses. The median (IQR) percentage of overlapping authors between the primary and secondary publications set was 64% (40–89%).

### 3.3. Predictors of having a secondary publication

The probability of having at least one secondary publication was 15% at 1 year, 33% at 2 years, 40% at 3 years, and 45% at 4 years after the primary article (Fig. 2). In univariable analyses, sample size, general medical journal, presence of preprimary secondary publication, and presence of a design article predicted secondary publications. The same characteristics, as well as whether the trial was industry funded, predicted more, recurrent secondary publications. In multivariable modeling, greater sample size, presence of preprimary secondary publication and presence of a design article predicted secondary publications. The first two characteristics predicted more, recurrent secondary publications (Table 3).

### 3.4. In-depth evaluation of types and results of secondary publications

The one secondary publication [25] testing the same hypothesis as the primary publication [29] was commissioned by the FDA to re-evaluate the key outcomes, trial definitions and event rates of the primary trial. Specifically, the primary trial [29] evaluated cardiovascular outcomes after addition of rosiglitazone to either metformin or sulfonylurea compared with the combination of the two, whereas the secondary reanalysis [25] was commissioned by the FDA to re-evaluate all deaths, suspected myocardial infarctions, and strokes blinded to treatment; derive end-of-follow-up dates; adjudicate suspected events by original trial definitions and by contemporary definitions under development by the FDA; and report event rates and time-to-event analyses by treatment group. The reanalysis did not produce a different interpretation as to which patients should be treated with the experimental intervention.

### Table 1. Study characteristics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Primary articles (n = 191),</th>
<th>Secondary articles (total sample, n = 475),</th>
<th>Secondary articles (restricted sample, n = 197)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 2009</td>
<td>—</td>
<td>28 (6)</td>
<td>13 (7)</td>
</tr>
<tr>
<td>2009</td>
<td>191 (100)</td>
<td>42 (9)</td>
<td>14 (7)</td>
</tr>
<tr>
<td>2010</td>
<td>—</td>
<td>81 (17)</td>
<td>42 (21)</td>
</tr>
<tr>
<td>2011</td>
<td>—</td>
<td>114 (24)</td>
<td>55 (28)</td>
</tr>
<tr>
<td>2012</td>
<td>—</td>
<td>107 (23)</td>
<td>41 (21)</td>
</tr>
<tr>
<td>2013</td>
<td>—</td>
<td>97 (20)</td>
<td>30 (15)</td>
</tr>
<tr>
<td>2014</td>
<td>—</td>
<td>6 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>General medical journal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>83 (43)</td>
<td>22 (5)</td>
<td>10 (5)</td>
</tr>
<tr>
<td>No</td>
<td>108 (57)</td>
<td>453 (95)</td>
<td>187 (95)</td>
</tr>
<tr>
<td>Country(ies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europeab</td>
<td>87 (43)</td>
<td>212 (45)</td>
<td>94 (48)</td>
</tr>
<tr>
<td>United States</td>
<td>81 (41)</td>
<td>360 (76)</td>
<td>102 (52)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>47 (24)</td>
<td>91 (19)</td>
<td>47 (24)</td>
</tr>
<tr>
<td>Canada</td>
<td>29 (15)</td>
<td>64 (13)</td>
<td>22 (11)</td>
</tr>
<tr>
<td>Oceania</td>
<td>22 (11)</td>
<td>30 (6)</td>
<td>13 (7)</td>
</tr>
<tr>
<td>Asia</td>
<td>16 (8)</td>
<td>60 (13)</td>
<td>16 (8)</td>
</tr>
<tr>
<td>Africa</td>
<td>8 (4)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>South America</td>
<td>5 (3)</td>
<td>18 (4)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>North Americaab</td>
<td>5 (3)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Industry funded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>82 (43)</td>
<td>190 (40)</td>
<td>77 (39)</td>
</tr>
<tr>
<td>No</td>
<td>109 (57)</td>
<td>285 (60)</td>
<td>120 (61)</td>
</tr>
</tbody>
</table>

a Not including the United Kingdom.
b Not including United States and Canada.
Of the 193 secondary publications testing different hypotheses compared to the primary article, 43 (22%) tested differences between subgroups that had not been addressed in the original article, of which 32 (74%) had at least one statistically significant result. Eighty-five (44%) publications assessed predictive factors associated with an outcome of interest, of which 72 (85%) found at least one statistically significant predictor. One hundred eighteen (61%) publications evaluated different outcomes than the original article, of which 95 (81%) found at least one statistically significant result. Seventy-one (37%) had differences in eligibility criteria, of which 52 (73%) found at least one statistically significant result. Twenty-one (10%) assessed different durations of follow-up; 12 (57%) had statistically significant results; 20 of 21 (95%) were concordant with the original primary trial findings (both significant n = 15, both nonsignificant n = 5, significant only in the primary n = 1).

Of the 193 secondary articles including analyses of different types, 176 (91%) had at least one analysis that found statistically significant results. Of the 17 that were not significant, 13 were concordant with the original primary trial findings (both nonsignificant n = 15, both statistically significant n = 0).

---

**Table 2. Description of trials with >10 identified secondary publications**

<table>
<thead>
<tr>
<th>Author year</th>
<th>Population</th>
<th>Interventions</th>
<th>Multisite</th>
<th>Sample size</th>
<th>Industry funded</th>
<th>Presence of a design article</th>
<th>Number of secondary publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andriole 2009</td>
<td>Male patients aged 54–74 years with history of prostate disease</td>
<td>Screening with serum prostate-specific antigen testing vs. usual care</td>
<td>Yes</td>
<td>76,693</td>
<td>Yes</td>
<td>Yes</td>
<td>35</td>
</tr>
<tr>
<td>Cummings 2009</td>
<td>Osteoporosis</td>
<td>Denosumab vs. placebo</td>
<td>Yes</td>
<td>7,868</td>
<td>Yes</td>
<td>No</td>
<td>16</td>
</tr>
<tr>
<td>Foster 2009</td>
<td>Overweight and obese adults with type 2 diabetes</td>
<td>Diabetes support and education vs. intensive lifestyle intervention</td>
<td>Yes</td>
<td>264</td>
<td>No</td>
<td>Yes</td>
<td>38</td>
</tr>
<tr>
<td>Mehran 2009</td>
<td>Acute myocardial infarction</td>
<td>Bivalirudin vs. heparin plus a glycoprotein IIb/IIIa inhibitors</td>
<td>Yes</td>
<td>3,602</td>
<td>Yes</td>
<td>Yes</td>
<td>51</td>
</tr>
<tr>
<td>Moss 2009</td>
<td>Heart failure</td>
<td>Cardiovascular resynchronization therapy plus an implantable cardioverter</td>
<td>Yes</td>
<td>1,820</td>
<td>Yes</td>
<td>Yes</td>
<td>34</td>
</tr>
<tr>
<td>O'Connor 2009</td>
<td>Medically stable outpatients with heart failure and reduced ejection fraction</td>
<td>Usual care plus aerobic exercise training, consisting of 36 supervised</td>
<td>Yes</td>
<td>2,331</td>
<td>No</td>
<td>Yes</td>
<td>39</td>
</tr>
<tr>
<td>Piper 2009</td>
<td>Smoking cessation</td>
<td>Nicotine lozenge, nicotine patch, sustained-release bupropion vs. nicotine</td>
<td>Yes</td>
<td>1,504</td>
<td>Yes</td>
<td>No</td>
<td>13</td>
</tr>
<tr>
<td>Wallentin 2009</td>
<td>Acute coronary syndromes</td>
<td>Ticagrelor vs. clopidogrel</td>
<td>Yes</td>
<td>18,624</td>
<td>Yes</td>
<td>Yes</td>
<td>31</td>
</tr>
</tbody>
</table>

---

![Fig. 2. Kaplan–Meier plot of time to having a first secondary publication.](image)
nonstatistically significant, 8 (47%) focused on adverse events where a nonsignificant result was actually favorable for the experimental intervention [30–36], 2 (12%) were cost-effectiveness analyses [37,38], and 7 (41%) pertained to other nonsignificant analyses, although two of them still considered that the hypothesis was still potentially important despite the negative result [39,40]. Appendix C at www.jclinepi.com provides raw data for the types of analyses conducted in each secondary article.

### 3.5. Specification of secondary analyses in trial protocols

From the 191 primary articles, we received 59 (30%) protocols from study authors. Of the 59 trials providing protocols, 28 (47%) published secondary analyses. Of these, 11 (39%) specified all secondary analyses in their protocol, 13 (46%) specified some but not all published secondary analyses in the protocol, and 4 (14%) did not specify any of their secondary analyses in the protocol.

### 4. Conclusions

We found that approximately half of the primary trials in high-impact journals have secondary publications appearing within a few years after the primary publication. A few trials produced a very large number of secondary publications (with eight trials [4% of primary articles] producing more than half of the secondary publications that were identified), but the most common scenario was to have 0, 1, or 2 secondary articles. A greater sample size, presence of preprimary secondary publications, and a previously published design article predicted more secondary publications. These are generally characteristics of larger pivotal trials, which tend to collect more data and thus have more opportunities for additional analyses and secondary publications in specialty journals [41].

In our evaluation, almost all secondary articles tested different hypotheses than the primary article, which is not surprising as it is rare for authors to reanalyze the same hypothesis [15]. Interestingly, almost all secondary articles found a statistical significance for some secondary analysis. Among the few nonsignificant ones, many were still favorable for the experimental intervention (e.g., documenting no excess in harms). In the current publishing environment, secondary articles seem to require at least one significant or favorable result to define a publishable unit [42,43].

Secondary publications may be useful if they convey important results that were not initially published in the primary article. On the other hand, readers may be confused by fragmentation of results of the same trial across many

---

**Table 3. Factors associated with having a secondary publication**

<table>
<thead>
<tr>
<th>Primary publication factor</th>
<th>First secondary publication (n = 197)</th>
<th>Recurrent secondary publications (n = 475)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariable analysis; HR (95% CI), P-value</td>
<td>Multivariable analysis; HR (95% CI), P-value</td>
</tr>
<tr>
<td>Number of primary outcomes</td>
<td>Reference</td>
<td>1.44 (0.90–2.29), 0.13</td>
</tr>
<tr>
<td></td>
<td>&gt; 1</td>
<td>1.00 (0.99–1.02), 0.99</td>
</tr>
<tr>
<td>Multisite</td>
<td>Yes</td>
<td>1.92 (1.48–2.51), &lt;0.001</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1.57 (1.09–2.55), 0.02</td>
</tr>
<tr>
<td>Sample size (per 10-fold increase)</td>
<td>Yes</td>
<td>1.67 (1.09–2.55), 0.02</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Reference</td>
</tr>
<tr>
<td>General medical journal</td>
<td>Yes</td>
<td>1.52 (0.99–2.31), 0.054</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Reference</td>
</tr>
<tr>
<td>Industry funded</td>
<td>Yes</td>
<td>4.63 (2.59–8.28), &lt;0.001</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Reference</td>
</tr>
<tr>
<td>Preprimary secondary publication</td>
<td>Yes</td>
<td>3.00 (1.76–5.12), &lt;0.001</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Reference</td>
</tr>
<tr>
<td>Design article</td>
<td>Yes</td>
<td>1.06 (0.88–1.28), 0.56</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Reference</td>
</tr>
<tr>
<td>Significance of at least one primary outcome</td>
<td>Yes</td>
<td>1.06 (0.88–1.28), 0.56</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Reference</td>
</tr>
</tbody>
</table>

**Abbreviations:** HR, hazard ratio; CI, confidence interval.
articles. For some pivotal trials, numerous secondary publications create publication factories. The main challenge is to which extent these significant findings are a result of spurious multiple testing. When results from multiple comparisons are presented across articles, the reader is unaware of how many hypotheses have been tested on trial data sets across multiple articles. Moreover, the multiplicity of these analyses is never corrected for, leaving readers unaware of the high likelihood of false positive results. Most of the secondary articles present subgroup or predictive analyses or evaluations of multiple outcomes, and this has been documented also in a previous empirical evaluation of a more limited sample \((n = 69)\) of secondary publications by Hopewell et al. [41]. Chance findings are notoriously common in claims of subgroup analyses [44–47]. Similar considerations apply when multiple predictors are assessed (effects may be spurious or inflated) [48,49], and when multiple outcomes are analyzed, especially when there is no clear prespecification of the outcomes in the protocols and selective reporting ensues [50–52]. One should be cautious when interpreting findings from secondary analyses and one would be more confident in the results if they were replicated in an independent trial or were consistent across multiple trials.

There are some limitations in our study. We excluded secondary articles that might have been performed by other teams using the raw data but without including any authors from the primary article. Given that data sharing from clinical trials has been uncommon until now and most analyses of the same trial data set are performed by the original authors [15], we doubt that many such articles were missed. We chose a random sample of primary articles from the top medical and specialty journals; these trials may be more prolific in generating secondary articles than trials published in lower impact journals. Thus, our sample likely produces an overestimation of secondary articles compared to the average trial. However, it offers appropriate estimates for a cohort of trials in influential journals that are more likely to affect clinical practice. For our in-depth analyses, we used a restricted subset of secondary articles when there were more than 3 secondary publications that had been published on the same primary article instead of the entire secondary publications set. Given that the characteristics are similar across the restricted and larger secondary publications set, this subset seems representative of the larger set.

In some occasions, the number of articles related to a randomized trial may be even larger than what we have identified. Some trials, especially those done by very large multi-center groups, may have published some articles that have no overlapping authors with those of the primary article, but may still belong to the same wider team where different authors rotate across different articles, sometimes without overlapping authorship. A few secondary analyses may also not cite the primary article at all; they may have no citation to the primary study, or they may cite a design article that has not been cited by the primary article or they may cite some other secondary publication(s) instead. An even more frequent occurrence may be the publication of additional meeting abstracts and articles, including editorials, and reviews. This creates an even bigger volume of secondary and related articles for some key trials with prolific authors. This volume may create both fragmentation and confusion for someone trying to assess the overall evidence.

We did not examine whether the secondary publications performed formal power calculations for the hypotheses that they evaluated. Nevertheless, the performance of multiple analyses scattered across multiple articles poses a multiplicity burden that is difficult to take properly into account in making power calculations or assessing the meaning of reported \(P\)-values that are typically unadjusted for these multiple comparisons. In contrast to secondary/sensitivity analyses that may be placed in a supplement of the original article, these additional analyses appeared in separate articles of their own. A full appraisal of the multiplicity burden would require knowing all the articles published from the same data set, something that currently is not readily visible when one comes across a secondary publication.

In all, the paradigm where each trial leads to a single article including all the relevant results has been abandoned by many trialists. The proliferation of articles from the same trial may help build CVs for investigators and may avoid the traditional drawback of trials that have commonly been called the deathbed of the assistant professor, because, in contrast to observational studies, they take a long time to complete but generate only a single article. However, this proliferation may also create unnecessary redundancy, introduce post hoc multiplicity that is not accounted for, and damage science as well confuse readers and clinicians. Therefore, we suggest that it would be optimal if indexing of secondary articles could be made in a way that all articles from a trial can be bundled together. Some trialists already maintain web sites where they list all the articles they have published from a trial data set and related work (e.g., the Women’s Health Initiative (WHI) trial [53]; Sleep AHEAD study [19]; the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening trial) [17]. Bundling together all articles from a given trial would help the reader understand the full breadth of the analyses that have been reported [54], consider whether multiplicity adjustments are appropriate, interpret cautiously the full spectrum of outcomes and results, use more complete and complementary information from the trial in systematic reviews and meta-analyses, and avoid confusion about whether separately published results represent the same or different studies. Public availability of trial protocols would also allow comparing the reported analyses against those that were presuppecified and see which analyses were post hoc. Such a systematic approach may make secondary publications more useful than confusing.
Acknowledgments

Authors’ contributions: S.E. and J.P.A.I. conceived the study design. S.E., L.M., M.K.e.D., Z.N.S., A.A., S.B., J.S., and N.S. screened, reviewed, and extracted data. S.E. and J.P.A.I. completed the statistical analyses. S.E. drafted the first draft of the article. All authors reviewed several drafts of the article and approval the final version of the article to be published.

Transparency declaration: S.E. and J.P.A.I. affirm that the article is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2016.05.016.

References


