A Methods Review of Posttrial Follow-up Studies of Cardiovascular Prevention
Finds Potential Biases in Estimating Legacy Effects

Lin Zhu*, Katy J.L. Bellb, Agnish Nayakc, Andrew Hayena

a School of Public Health, Faculty of Health, University of Technology Sydney, Sydney, New South Wales, Australia
b School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, New South Wales, Australia
c Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia

* Corresponding author:
School of Public Health, Faculty of Health, University of Technology Sydney, Sydney, New South Wales, Australia
Tel.: +61 295145014;
E-mail address: Lin.Zhu@uts.edu.au
Abstract

Objectives: To assess the methods used, and potential for bias, in post-trial studies of cardiovascular disease (CVD) where legacy effects may be estimated.

Study Design and Setting: We undertook a methods review of post-trial studies after randomized controlled trials of interventions to prevent CVD. For each included article, we extracted information on important aspects of the design and analysis of the study, and on the reporting of legacy effects.

Results: Of 2622 retrieved articles, 46 were included in the review: 13 on blood glucose control, 13 on blood pressure control and 20 on blood lipids control. The median duration for the RCT and post-trial follow-up studies were 5.0 and 5.7 years respectively. At least 80% of initial RCT participants were enrolled in the post-trial study in 32 of the reports. Most reports used both linkage to routine administrative data sources and active data collection for the post-trial study. Of the 46 included articles, the authors assessed and reported post-trial covariate balance in 29, and made statistical adjustments in the analysis for potential confounding in 25. Post-trial results were reported separately to overall results (from time of randomization) in 21 articles. Legacy effects were claimed in 19 reports, of which 16 could be justified on the basis of the post-trial results.

Conclusion: Post-trial studies may provide valuable information for investigating legacy effects, but better reporting of results is needed to realize their full potential. Robust methods of data collection and analysis may address the risk of selection and confounding biases in post-trial studies.

Keywords: Legacy effects; Randomized controlled trial; Post-trial follow-up; Method Review
What is new?

Key findings

Fifty-four percent of post-trial studies reported only analyses that combined initial trial and post-trial follow-up data. This fails to make full use of the data and may lead to incorrect conclusions on legacy effects.

63% of post-trial studies compared between-group-difference of covariates for the participants enrolled in the post-trial period, and fifty-four percent of studies made adjustment in their analysis. Most comparisons and adjustments were made using baseline measurements of covariates at start of the RCT, rather than the more appropriate time point at the start of the post-trial study.

What this adds to what was known?

This review has added to current evidence that some post-trial studies use inappropriate methods of analysis, and that there is a need to improve the reporting of legacy effects.

What is the implication and what should change now?

Researchers should be aware of the potential for confounding and selection bias when designing post-trial studies. By using appropriate methods of data collection and analysis, they may minimise bias in estimates of long term and legacy treatment effects.
Introduction

The randomised controlled trial is the ideal study design for generating evidence on the effectiveness of clinical interventions. However, because trials are resource intensive to run, they commonly include relatively short follow-up periods and may fail to capture long-term effects of the intervention. A relatively low-cost option that trialists may choose to extend follow-up is to undertake a post-trial cohort study of surviving participants after the end of the active trial period [1]. The term “legacy effect” is often used in the context of such studies, which describes the effects of an intervention that are only observed after the end of trial and are not the direct effects observed during the trial period itself.

The finding of a legacy effect may provide support for earlier initiation at a younger age (or potentially cessation at a younger age) of the intervention under study. Much of the clinical interest in legacy effects has been in drug treatments for cardiovascular disease prevention. An increasing number of studies have reported legacy effects for cardiovascular prevention drug treatments on the basis of results from post-trial follow-up studies after randomized controlled trials [2–4]. Several systematic reviews have also investigated legacy effects [5–8].

Post-trial analyses, which assess the treatment effects beyond the end of the trial, can provide information for the management of patients [9]. However, there are analytic challenges for these studies that arise from comparing groups that, unlike at the start of the trial period, are no longer at equivalent risk. For instance, at the end of Diabetes Control and Complications Trial (DCCT), the conventional and intensive treatment randomized groups differed on several established CVD risk factors (e.g., BMI and triglycerides) and in levels of the surrogate outcome of asymptomatic microvascular disease (microalbuminuria and albuminuria)[10]. Other sources of bias include continued differential use of medication between study arms, treatment confounder feedback, differential loss to follow-up and differences in time-dependent covariates across the original randomized groups [11]. Analysis that fails to consider these issues may lead to biased results. In addition, the results of trial and post-trial
analyses may be mis-interpreted. Analyses combining both the initial trial period and
the post-trial follow-up period (long-term effect) have often been incorrectly
interpreted as evidence of a legacy effect, which is better assessed on the basis of
separate post-trial analysis [12]. An incorrect assumption of a beneficial legacy effect
may inappropriately influence clinicians to prescribe interventions more
aggressively at a younger age, or similarly influence clinical guidelines
recommendations to the same effect.

Although post-trial studies have become more common, limited attention has been
paid to the design and analysis of such studies. In this article, we conduct a review of
the methods used in post-trial studies after RCTs of drug interventions to prevent
cardiovascular disease. Our objectives were to assess the design characteristics,
methods of analysis used, and the quality of reporting of legacy effects.

Methods
We registered the protocol for this review on the international prospective register
of systematic reviews (PROSPERO, CRD42017063969). We included all reports on
follow-up after randomized, controlled trials in adults (age >18 years) of drug
interventions to lower blood glucose, blood pressure and cholesterol (including
placebo-controlled studies or lower vs higher targets for treatment).

Search Strategy
We searched the following electronic bibliographic databases (to 31st December
2019): MEDLINE, EMBASE and The Cochrane Library (Cochrane Database of
Systematic Reviews, Cochrane Central Register of Controlled Trials, Cochrane
Methodology Register). Other sources included reference lists of included papers,
hand searching key journals, conference proceedings, forward citation searching of
included studies. The search strategy is provided in supplementary materials.

Study Selection & Data Extraction
Detailed inclusion and exclusion criteria for study selection are listed in Table 1. Two
reviewers (LZ and AN) checked the titles and abstracts of all citations identified through the database searches and forward citation search. We obtained full text if either reviewer judged the article as being potentially relevant. The same two authors then independently checked all the full text articles for eligibility, resolving disagreements through discussion with two further reviewers (KB and AH).

**Table 1. Inclusion and exclusion criteria for study selection**

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Post-trial reports of randomized controlled trials evaluating intensive vs standard blood pressure control or lipid control or glucose control in diabetics (including placebo-controlled studies or lower vs higher targets for treatment).</td>
<td>● Clinical cardiovascular disease outcomes (fatal or non-fatal) during post-trial follow-up period not reported.</td>
</tr>
<tr>
<td>● Adults (≥ 18 years)</td>
<td>● Post-trial follow-up period less than 12 months</td>
</tr>
<tr>
<td>● The randomized controlled trial was ≥12 months duration</td>
<td>● Data not reported separately for each randomized group</td>
</tr>
<tr>
<td>● Studies in any setting (primary or secondary CVD prevention)</td>
<td></td>
</tr>
</tbody>
</table>

**Data extraction and Assessment**

LZ extracted data using standardized forms, and the results were reviewed by AH and KB. In addition to descriptive information of both the randomized controlled trial and post-trial follow-up, we extracted the following data on the design and analysis of the post-trial study, and on the reporting of legacy effect [13]:

(i) Method of data collection: The methods of data collection were divided into three categories: 1) active follow-up, such as clinic visits, laboratory examinations, questionnaires or telephone interviews; 2) data linkage, where follow-up data were obtained by linkage to administrative datasets such as hospital records, death certificates or population registries and 3) a combination of 1) and 2).

(ii) Information collected in the post-trial period, including medication use; relevant
surrogate outcomes (e.g. blood glucose, blood pressure, LDL cholesterol, total cholesterol); and other cardiovascular risk factors.

(iii) Assessment of between-group-difference in relevant covariates. We extracted information on comparisons of covariates between the trial arms of participants and on whether the comparison was made based on the covariates measured at baseline or during the study.

(iv) The primary statistical method used in the post-trial analysis.

(v) Statistical adjustment in the post-trial analysis. We extracted information on whether the analysis accounted for relevant covariates and whether the adjustment was made based on baseline data or updated data during the post-trial study.

(vi) Reporting of results: post-trial result only; trial and post-trial combined (overall result only); both post-trial result and overall result.

(vii) Types of effects claimed in the abstract: legacy effect, long-term effect, both legacy and long-term effects.

(viii) Justification of legacy effect: whether the claims on legacy effect were justified by reporting a separate post-trial result.

Results

Results of the search

Figure 1 shows the flow diagram of the post-trial study selection process. Of the 2622 records we identified, 2524 were ineligible based on screening of abstracts, and 98 full-text articles were assessed for eligibility. Of these 98 articles, we assessed 44 as eligible, and we identified two further articles through searches of references and forward citations. In total, we included 46 articles from 37 unique follow-up studies. Detailed information on included studies is provided in the supplementary materials Table S1.
Figure 1. Flow diagram of the post-trial study selection process

* 31 studies each had one corresponding report; 3 studies each had 2 corresponding reports (ADVANCE-ON Blood Glucose study, Veteran Affairs Diabetes Trial, Systolic Hypertension in the Elderly Program) and 3 studies each had 3 corresponding reports (Epidemiology of Diabetes Interventions and Complications study, Anglo-Scandinavian Cardiac Outcomes Trial, West of Scotland coronary prevention Study).

Characteristics of the RCTs and post-trial follow-up

We summarize the characteristics of the RCTs and post-trial follow-up in Table 2. Of the 46 articles, 13 (28%) investigated blood glucose control, 13 (28%) investigated blood pressure control and 20 (44%) investigated lipid control. About one third of these interventions were for primary prevention. The median duration for RCT and post-trial follow-up period were 5.0 and 5.7 years, respectively.

Thirty percent of RCTs recruited fewer than 2000 participants, while 40% had more than 5000 participants. Most articles reported a between group difference in levels of
surrogate outcomes (lipids, blood pressure, blood glucose) in the trial period, and about 60% articles reported a significant primary endpoint for trial. Although in all studies the investigated treatment was recommended to all subjects (on the basis of the RCT results), it was actively provided to all subjects in 5 post-trial studies while the other post-trial studies were purely observational. Thirty-four (74%) articles used the same primary outcome for the post-trial follow-up as the initial RCT.

Table 2. Characteristics of the RCTs and post-trial follow-up

<table>
<thead>
<tr>
<th>Period</th>
<th>Characteristics</th>
<th>No. of Reports (%, N=46)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of Intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood Glucose</td>
<td></td>
<td>13 (28)</td>
</tr>
<tr>
<td>Blood Pressure</td>
<td></td>
<td>13 (28)</td>
</tr>
<tr>
<td>Lipid</td>
<td></td>
<td>20 (44)</td>
</tr>
<tr>
<td><strong>Type of Prevention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary Prevention</td>
<td></td>
<td>15 (33)</td>
</tr>
<tr>
<td>Secondary Prevention</td>
<td></td>
<td>31 (67)</td>
</tr>
<tr>
<td><strong>Length of Trial Follow-up</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (Mean)</td>
<td></td>
<td>5.0 (4.9)</td>
</tr>
<tr>
<td>&lt;4</td>
<td></td>
<td>11 (24)</td>
</tr>
<tr>
<td>4-8</td>
<td></td>
<td>33 (72)</td>
</tr>
<tr>
<td>≥8</td>
<td></td>
<td>2 (4)</td>
</tr>
<tr>
<td><strong>No. of subject initial RCT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (Mean)</td>
<td></td>
<td>4446 (5236)</td>
</tr>
<tr>
<td>&lt;500</td>
<td></td>
<td>2 (4)</td>
</tr>
<tr>
<td>500-2000</td>
<td></td>
<td>12 (26)</td>
</tr>
<tr>
<td>2000-5000</td>
<td></td>
<td>15 (30)</td>
</tr>
<tr>
<td>5000-10000</td>
<td></td>
<td>8 (20)</td>
</tr>
<tr>
<td>≥10000</td>
<td></td>
<td>9 (20)</td>
</tr>
<tr>
<td><strong>Mean age of subject</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (Mean)</td>
<td></td>
<td>60 (58)</td>
</tr>
<tr>
<td>&lt;45</td>
<td></td>
<td>5 (11)</td>
</tr>
<tr>
<td>45-65</td>
<td></td>
<td>30 (65)</td>
</tr>
<tr>
<td>≥65</td>
<td></td>
<td>11 (24)</td>
</tr>
<tr>
<td><strong>Between-Group Different Surrogate</strong></td>
<td></td>
<td>43 (93)</td>
</tr>
<tr>
<td>Post-trial follow-up</td>
<td>Levels Reported</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statistically significant primary end point</td>
<td>27 (59)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Length of Post-trial Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median (Mean)</td>
</tr>
<tr>
<td></td>
<td>&lt;5</td>
</tr>
<tr>
<td></td>
<td>5-10</td>
</tr>
<tr>
<td></td>
<td>≥10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Percentage of subjects enrolled in Post-trial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median (Mean)</td>
</tr>
<tr>
<td></td>
<td>&gt;90%</td>
</tr>
<tr>
<td></td>
<td>80%-90%</td>
</tr>
<tr>
<td></td>
<td>≤80%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Between-group Difference in post-trial study enrollment*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;2%</td>
</tr>
<tr>
<td></td>
<td>2%-5%</td>
</tr>
<tr>
<td></td>
<td>≥5%</td>
</tr>
<tr>
<td></td>
<td>N.A.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Observational Post-trial Follow-up Study†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41 (89)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Primary outcome same as the RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>34 (74)</td>
</tr>
</tbody>
</table>

*Difference in enrollment of post-trial follow up between those initially enrolled in active and control groups
†Examining if the post-trial follow-up is observational or the post-trial treatment were provided by the investigators

Figure 2 summarizes the design and analysis of post-trial studies. At least 80% of the trial participants were enrolled in the post-trial follow-up study in 32 (70%) of the 46 articles. The percentage of participants enrolled in post-trial follow-up was similar between those who were initially in the intervention and in the placebo groups in 25 articles (difference less than 2%), while some reports showed larger differences. For example, in the Steno-2 study, the percentages of enrollment for the active and placebo group were 84% and 79% respectively. Among the 46 reports, 14 (30%) collected data by active follow-up; 12 (26%) obtained post-trial information by data linkage, and the remaining 20 (44%) reports used both methods. Data on medication
use in the post-trial period was collected by most reports (32, 70%). Surrogate outcomes were reported in 30 articles (65%), while information on other risk factors was available in 21 reports (45%).

The between-group-differences in covariates were assessed in 29 articles (63%), but in 22 (48%) of reports this was based on baseline measurements at start of the RCT rather than the post-trial study. Information on the between-group-difference with regards to covariates, surrogate outcomes and medication use in the post-trial period are provided in supplementary materials Table S2. Cox proportional hazard was the most frequently used method in the post-trial analysis (37, 80%). Twenty-five reports (54%) made statistical adjustments for covariates in their analysis. In addition to time invariant demographic factors such as age, gender and race, eight reports also adjusted for time-varying covariates. For example, in the Epidemiology of Diabetes Interventions and Complications (EDIC) study, updated mean glycosylated hemoglobin value, the development of renal disease, microalbuminuria and albuminuria were included as time-dependent covariates in analysis [10,14].

19 (41%) articles reported results for both post-trial and the whole study period, while 2 (4%) only reported the results of the post-trial period and 25 (54%) articles only reported an overall result. Long-term effects were claimed in the abstract of 27 (59%) articles. Legacy effects were claimed in the abstracts of 19 reports, of which 16 provided post-trial results that justified the claim.
**Figure 2.** The design, analysis and reporting of the post-trial studies

**Discussion:**
In this methods review of post-trial studies for a range of cardiovascular interventions, we found potential for biased estimates of legacy effects from the study designs and methods of analysis most commonly used. We also found important
deficiencies in the reporting of these studies, and in particular that a separate post-
trial analysis was infrequently reported.

We found that post-trial studies are often missing information on important
covariates that would allow assessment and adjustment for potential selection bias
and confounding [1,15]. About one third of the articles did not report the post-trial
medication use or measurements of surrogate outcomes (such as BP, HbA1c, lipids),
and the information on the other potential confounders was not collected in more
than half of the reports. Where initial RCTs are very large, it may be prohibitively
expensive to collect the data of all the patients [16]. Random sampling may be a
solution for such studies. For instance, of 8,494 participants enrolled in the post-trial
study of the Action in Diabetes and Vascular Disease trial (ADVANCE-ON), a random
subset of 2,000 patients, balanced across regions and across the prior randomized
treatment arms, were invited to undergo laboratory examinations to determine
whether in-trial differences in surrogate outcomes persisted in post-trial [17].

Even where sufficient data were collected that could allow for assessment and
adjustment for bias in analysis, we found that most post-trial studies failed to do so.
The main challenge of post-trial analyses is that there is no longer a randomized
comparison of intervention vs control, and the study design is that of a cohort study.
Although 46% studies collected information on post-trial covariates, only about 15%
studies assessed the balance of post-trial covariates and made corresponding
adjustment in their analysis. As it is hard to determine to what extent the risk of CVD
remains equivalent, we suggest that both unadjusted and adjusted results should be
reported to ensure the robustness of findings. In addition, methods used to deal with
time-dependent confounding and selection biases in observational studies, such as
causal inference approaches (G-methods), could also be applied to the post-trial
analysis [18].

Our review also highlights the need for improved reporting of treatment effects in
post-trial studies. We found that most studies only focused on long-term treatment
effects [19,20]. Among the articles claiming legacy effect, about one quarter failed to report a separate post-trial result to justify the evidence. As both direct effects and legacy effects could contribute to the long-term effects of the intervention, long-term effects alone do not provide proof of legacy effects. For example, in the Scandinavian Simvastatin Survival Study, the investigators found simvastatin treatment for 5 years in trial was associated with a survival benefit over 10 years. The hazard ratio and corresponding confidence interval reported for trial and whole follow-up period were 0.70 (0.58–0.84) and 0.85 (0.74–0.97), respectively. However, the risk calculated by the separate post-trial analysis was 1.03 (0.86–1.24) [21]. While a sustained survival benefit was found after the end of the trial, it was apparently due to the direct effect during the trial, rather than a legacy effect emerging in the post-trial follow-up.

Our review has some limitations. We only focused on the post-trial studies of RCTs evaluating cardiovascular interventions. In previous reviews of long-term follow-up of randomized trial participants, other common types of interventions included surgery, cancer screening and behavioral change interventions[1,15]. The extent to which the issues we found applies to these other interventions and settings is unknown. We also did not search for non-English language studies, nor the gray literature (including unpublished studies). However, it seems unlikely that we missed important methods in our search.

Conclusion
As post-trial studies are becoming increasingly common, it is important for readers to be aware of important methodological issues related to the estimation and interpretation of legacy effects. Trialists aiming to investigate legacy effects need to ensure appropriate study design and methods of analysis are used. In particular, trialists should design and conduct post-trial studies using methods of data collection and analysis that allow potential confounding and selection bias to be addressed.

Declaration of interest
The author declares no conflict of interest.

**Source of funding:**

Lin Zhu was supported by an Australian Government Research Training Program Scholarship at the University of Technology Sydney.

**Author contributions**

Lin Zhu: Conceptualization, Investigation, Methodology, Writing - original draft. Katy J.L. Bell: Conceptualization, Investigation, Methodology, Supervision, Writing - review & editing. Agnish Nayak: Investigation, review & editing. Andrew Hayen: Conceptualization, Investigation, Methodology, Supervision, Writing - review & editing.

**References:**


