



## Methods

# Comparing methods to address bias in observational data: statin use and cardiovascular events in a US cohort

Paulina Kaiser,<sup>1\*</sup> Alice M Arnold,<sup>2</sup> David Benkeser,<sup>2</sup>  
Adina Zeki Al Hazzouri,<sup>3</sup> Calvin H Hirsch,<sup>4</sup> Bruce M Psaty<sup>5</sup> and  
Michelle C Odden<sup>1</sup>

<sup>1</sup>College of Public Health & Human Sciences, Oregon State University, Corvallis, OR, USA, <sup>2</sup>Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA, USA, <sup>3</sup>Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL, USA, <sup>4</sup>Center for Healthcare Policy and Research, University of California Davis, Davis, CA, USA and <sup>5</sup>Cardiovascular Health Research Unit, Department of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, Washington and Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA

\*Corresponding author. College of Public Health & Human Sciences, 155 Milam Hall, Corvallis, OR 97331, USA.  
E-mail: paulina.kaiser@oregonstate.edu

Editorial decision 31 July 2017; Accepted 8 August 2017

## Abstract

**Background:** The theoretical conditions under which causal estimates can be derived from observational data are challenging to achieve in the real world. Applied examples can help elucidate the practical limitations of methods to estimate randomized-controlled trial effects from observational data.

**Methods:** We used six methods with varying design and analytic features to compare the 5-year risk of incident myocardial infarction among statin users and non-users, and used non-cardiovascular mortality as a negative control outcome. Design features included restriction to a statin-eligible population and new users only; analytic features included multivariable adjustment and propensity score matching.

**Results:** We used data from 5294 participants in the Cardiovascular Health Study from 1989 to 2004. For non-cardiovascular mortality, most methods produced protective estimates with confidence intervals that crossed the null. The hazard ratio (HR) was 0.92, 95% confidence interval: 0.58, 1.46 using propensity score matching among eligible new users. For myocardial infarction, all estimates were strongly protective; the propensity score-matched analysis among eligible new users resulted in a HR of 0.55 (0.29, 1.05)—a much stronger association than observed in randomized controlled trials.

**Conclusions:** In designs that compare active treatment with non-treated participants to evaluate effectiveness, methods to address bias in observational data may be limited in real-world settings by residual bias.

**Key words:** Bias, observational studies, statins

### Key Messages

- There is a need for better understanding of the practical limitations of methods to estimate randomized-controlled trial (RCT) effects from observational data.
- We compared six methods to estimate the effect of statin use on myocardial infarction and non-cardiovascular mortality using data from the Cardiovascular Health Study.
- The methods used had a modest effect on estimated associations, particularly for myocardial infarction, likely due to residual healthy user bias and limited precision.
- All methods produced stronger protective estimates of the effect of statins on myocardial infarction than comparable RCTs.

## Introduction

Drawing causal inferences from observational data often relies on estimating a counterfactual contrast for the exposure or treatment of interest. Substantial literature has considered statistical methods to estimate this counterfactual contrast without the benefit of randomization.<sup>1–3</sup> These methods have focused on recreating the study design conditions of a randomized-controlled trial (RCT), often including restriction of the study population to those eligible for the treatment<sup>4–9</sup> and restriction to new users of a treatment to allow adjustment for pre-treatment levels of covariates.<sup>10</sup> Additionally, analytic features such as adjustment for covariates and the use of propensity scores to quantify the likelihood of receiving treatment have been used to address differences between users and non-users.<sup>11,12</sup>

The combination of these design and analytic conditions has been successful in replicating results of trials, e.g. in an observational study of hormone replacement therapy on coronary heart disease (CHD) in the Nurses' Health Study<sup>5</sup> and of statins and CHD in a large database of electronic health records.<sup>6</sup> Given the known limitations of randomized trials,<sup>13,14</sup> these methods may be useful for evaluating potential benefits and harms of recently FDA-approved medications in post-marketing surveillance. However, the conditions under which causal estimates can be gleaned from observational data have not been clearly elucidated. The utility of these methods may differ depending on the outcome of interest; outcomes that the treatment is intended to effect (e.g. effectiveness outcomes) may be subject to more bias in observational data than outcomes that the treatment is not intended to affect (e.g. safety outcomes or adverse effects).<sup>15,16</sup> Furthermore, these methods have not been well evaluated in moderately sized cohorts, in which restriction-based methods to reduce bias may have a critical impact on sample size and, subsequently, precision.

We used data from the Cardiovascular Health Study (CHS) to apply a combination of design and analytic

strategies to a well-characterized cohort of 5888 older adults, to assess the utility of these methods in addressing bias related to statin use. We chose two outcomes: incident myocardial infarction (MI), as an effectiveness outcome; and non-cardiovascular (CVD) mortality, as a negative control outcome (based on biologic plausibility as well as prior studies of statin use).<sup>17–19</sup> Statins provide a useful setting for investigating issues of bias because they are a common preventive treatment, and the broad array of beneficial outcomes associated with statin use have raised questions about the potential pervasive influence of healthy user bias.<sup>20–22</sup> Healthy user bias results when a treatment is associated with other health-promoting characteristics, such as access to care and proactive health behaviours, leading to overly protective estimates of effectiveness. This is in contrast to confounding by indication, or the observation that less healthy individuals are more likely to be prescribed treatment, which would bias the estimate of effect in the harmful direction.<sup>23</sup>

CHS is an advantageous cohort for this analysis because of the large number of new statin users; study enrolment began just 2 years after FDA approval of the first statin. A prior publication from CHS suggested a strong protective association between statin use and incident cardiovascular events.<sup>24</sup> Subsequently, a large RCT (PROSPER) among high-risk adults aged 70–82 years suggested that the effect of statins was much smaller.<sup>25</sup> These disparate results suggest the possibility of healthy user bias or early adopter bias—an exaggerated form of healthy user bias unique to the time period immediately after release of a novel therapy. We aimed to evaluate whether a combination of design and analytic approaches using observational data would produce estimates consistent with RCTs.

## Methods

### Data source

CHS is a longitudinal cohort of 5888 adults aged 65 and over recruited from Medicare eligibility lists in four sites

(Sacramento County, CA; Washington County, MD; Forsyth County, NC; and Allegheny County, PA). Recruitment of the initial cohort of 5201 began in 1989, with an additional 687 African-American adults recruited in 1992.<sup>26</sup> Exclusion criteria included institutionalization, active cancer treatment or expectation of moving from the area within 2 years. Participants provided informed consent, and institutional review boards at each site approved the study.

Participants were contacted semi-annually, alternating between a telephone interview and an annual clinic examination through 1999, and continuing by telephone interview afterwards. Additional information was collected from medical records and interviews with surviving participants or proxies.

### Study design

We compared six models with various design and analytic features: (i) crude analysis, (ii) restriction to those eligible for statins, (iii) multivariable adjusted among eligible population, (iv) restriction to eligible new users, (v) multivariable adjusted among eligible new users and (vi) propensity score (PS)-matched among eligible new users. We applied each method to the association between statins and two outcomes of interest: non-CVD mortality, as a negative control outcome, and incident MI. We considered each study visit as a 'trial', and outcomes in the subsequent 5 years were evaluated in relation to statin status at the beginning of that trial (akin to an intent-to-treat approach of a 5-year randomized trial). Trials were pooled into a single analysis, and standard errors were adjusted appropriately to account for the fact that participants could contribute to more than one trial. This approach has been described previously.<sup>6,27,28</sup>

Participants were considered eligible for statins if they met any of the following criteria, based on guidelines informed by the National Cholesterol Education Program's 1993 Adult Treatment Panel II guidelines<sup>29</sup> and common practices in the early 1990s: total cholesterol above 240 mg/dL or low-density lipoprotein (LDL)-cholesterol above 190 mg/dL; LDL-cholesterol above 130 mg/dL with prevalent CHD; or LDL-cholesterol above 130 mg/dL with one or more risk factors (current smoker, diabetic, hypertensive, HDL-cholesterol less than 35 mg/dL, history of transient ischemic attack, history of stroke or sibling history of early MI, with HDL-cholesterol >60 mg/dL negating one risk factor). Eligibility was calculated at every visit, based on covariate values from the previous visit. For statin users, eligibility status at time of initiation was carried forward for the duration of statin use (e.g. to avoid eligibility status changing as a result of lowered LDL from taking statins).

### Measurement of key variables

#### Exposure: statin use

At each visit, medication use during the past 2 weeks was determined by inspection of prescription bottles.<sup>30</sup> Statin initiation was marked at the first study visit where statins were used, if statins were not used at the previous visit (to approximate a 1-year washout period, though precise time of initiation is unknown). Statin users whose statin status at the previous visit was missing were classified as statin users but not statin initiators. Statin use was not imputed; 2.9% of observations were missing statin use.

#### Outcomes: non-CVD mortality and incident MI

We included outcomes that occurred within 5 years after statin measurement. Cause of death (CVD or non-CVD) was adjudicated by a Morbidity and Mortality Committee.<sup>31</sup> The most common causes of non-CVD mortality were dementia, cancer and pneumonia. MI was also adjudicated by committee, and was defined on the basis of cardiac enzyme levels, chest pain and serial electrocardiographic changes.<sup>31</sup>

#### Additional covariates

Baseline age was modelled linearly. CHD status (defined as MI, angina, coronary bypass or angioplasty) was determined at each visit, by comparing the date of the visit with incidence dates of CHD as established by committee. Total cholesterol was collected every year through 1999 except for 1990, 1991 and 1995; LDL-cholesterol and HDL-cholesterol were only available in 1989 and 1992. Hypertension status [as normal, borderline (systolic blood pressure 130–9 mmHg or diastolic blood pressure 80–90 mmHg) or hypertensive (systolic blood pressure above 140 mmHg, diastolic blood pressure above 90 mmHg or use of anti-hypertensive medication as determined by medication inventory interview and a self-report of a history of high blood pressure)] was assessed every year except 1995. Smoking status (current, former or never) was assessed every year. Anti-hypertensive medication use was documented every year through medication inventory interview. Diabetes (use of insulin or oral hypoglycaemic agents or fasting glucose  $\geq 126$  mg/dL) was available in 1989, 1992 and 1996. Total cardiovascular risk points was calculated by giving one point each for smoking, diabetes, hypertension, low LDL, history of transient ischemic attack or stroke, claudication and family history of early MI (brother before age 55 or sister before age 65), with one point deducted for high LDL.

General health status (dichotomized to excellent or very good health vs good, fair or poor) was assessed every year except 1991. Education was measured at baseline and dichotomized for analysis as completion of 4 or more years

of college. Income was measured at baseline and dichotomized for analysis at \$35 000 per year or more. Additional insurance status (private vs other) was collected in 1993, 1994, 1996, 1997 and 1998. Living arrangement was dichotomized to living with spouse or other, and was available in 1989, 1992 and 1998.

Variables that were not measured annually were carried forward from the last available visit, for no more than 5 years. Missing observations were replaced with a non-missing value from the previous visit. Valid response categories for missing values were created for living arrangement (missing for 23.2% of observations) and additional insurance status (22.1% missing). With these adjustments, 2.0% of observations had any missing information.

## Statistical analysis

### Method 1: prevalent users, unadjusted

Using an unadjusted proportional hazards model, we compared all prevalent statin users to all non-users in the pseudo-trial framework as a naïve approach.

### Method 2: eligible prevalent users, unadjusted

We fit an unadjusted proportional hazards model among an analytic sample restricted to those who met the eligibility criteria for statins as described above.

### Method 3: eligible prevalent users, adjusted

We restricted the analytic sample to those who met the eligibility criteria described above, and fit a proportional hazards model adjusted for visit year, baseline age, sex, race, CHD status, total cholesterol, LDL-cholesterol, HDL-cholesterol, diabetes, oral hypoglycaemic medication, hypertension, anti-hypertensive medication, smoking status, total cardiovascular risk points, general health status, education, insurance status and living arrangement.

### Method 4: eligible new users, unadjusted

We restricted the analytic sample to those who met the eligibility criteria above and were not taking statins at the previous visit, and fit an unadjusted proportional hazards model. This model estimates the marginal intent-to-treat hazard ratio (HR) in a population that is eligible for statins according to the National Cholesterol Education Program's 1993 Adult Treatment Panel II guidelines.

### Method 5: eligible new users, adjusted

The analytic sample was restricted to those who met the eligibility criteria above and were not taking statins at the previous visit; we fit a proportional hazards model adjusted for visit year, baseline age, sex, race, CHD status, total cholesterol, LDL-cholesterol, HDL-cholesterol,

diabetes, oral hypoglycaemic medication, hypertension, anti-hypertensive medication, smoking status, total cardiovascular risk points, general health status, education, insurance status and living arrangement. This model estimates the intent-to-treat HR conditional on the above covariates among a statin-eligible population.

### Method 6: eligible new users, PS-matched

The analytic sample was restricted to those who met the eligibility criteria above, were not taking statins at the previous visit and were matched on their propensity for initiating statins. The treatment model included the same covariates used in the multivariable adjusted models above (visit year, baseline age, sex, race, CHD status, total cholesterol, LDL-cholesterol, HDL-cholesterol, diabetes, oral hypoglycaemic medication, hypertension, anti-hypertensive medication, smoking status, total cardiovascular risk points, general health status, education, insurance status and living arrangement). Matching was done by identifying the three non-initiators with propensity scores closest to the propensity score for each statin initiator, within a calliper of 0.001 (nearest neighbour matching with replacement). Of the 222 eligible initiators, 214 were matched successfully. The treatment model produced good covariate balance between matched initiators and non-initiators, as visualized by the standardized percentage bias ([Supplementary Figure 1](#), available as [Supplementary Data](#) at *IJE* online).<sup>32</sup> The analytic model estimates the intent-to-treat HR comparing statin initiators to non-initiators within a population matched on the predicted probability of treatment.

To evaluate the potential for shifting biases as statin use became more prevalent, we compared early statin initiators (those who started statins in 1990–95) to late statin initiators (those who initiated in 1996–98). We also investigated effect modification by age (age 65–69 vs 70+ at baseline), sex and propensity for statin initiation, given that statins have been demonstrated to be more effective in preventing MIs in secondary prevention and/or among those at highest cardiovascular risk.<sup>25</sup>

In addition, we replicated Methods 1–6 above as logistic models, in order to apply probability weights. As a seventh method, we used the propensity score treatment model described above to calculate inverse probability of treatment weights. To address a modest violation of the positivity assumption, we excluded the non-initiators with a predicted probability of treatment above or below the range of the initiators in this model. We also calculated inverse probability of attrition weights to investigate the impact of selective attrition and applied them to all seven methods (see [Supplementary Table 2](#), available as [Supplementary Data](#) at *IJE* online).

## Results

The current analysis excluded 32 participants with no data on statin use and 562 participants with a history of MI at baseline, leaving 5294 in the analytic sample. A total of 1625 non-CVD deaths and 757 incident MIs were observed within 5 years of statin measurement (120 non-CVD deaths and 41 incident MIs among statin users). A total of 697 people initiated statins during follow-up, including 66 who initiated statins twice and 3 who initiated statins three times. Statin initiation became more common during follow-up, with 0.7% of participants initiating at the first follow-up visit and 2.2% of participants initiating at the last follow-up visit (Supplementary Table 1, available as Supplementary Data at *IJE* online). Adherence to statins was good, with 73.4% of statin initiators still taking statins after one year and 68.7% still taking statins after 2 years (Supplementary Table 1, available as Supplementary Data at *IJE* online).

Statin initiators had higher LDL-cholesterol levels and body mass indices (BMIs) than non-users, and were more likely to have CHD, diabetes and be taking oral hypoglycaemic medications (Table 1). However, fewer than two-thirds of statin initiators were eligible for statins based on covariate values at the visit before initiation (Table 1).

Most point estimates of the association between statin use and 5-year non-CVD mortality were protective, with

confidence intervals that included the null (Figure 1). The unadjusted prevalent user model produced the estimate furthest from the null [HR 0.69, 95% confidence interval (CI): 0.55, 0.85], whereas restriction to those eligible for statins and new users attenuated the estimate. The propensity score-matched analysis produced a largely null estimate [HR 0.92 (0.58, 1.46)], but also the largest CI.

Estimates of the association between statin use and incident MI were broadly similar across all methods, with overlapping confidence intervals (Figure 2). Crude models showed a strong protective association between statin use and 5-year incident MI [HR 0.54 (0.37–0.78)]. Additional restrictions and multivariable adjustment resulted in slightly stronger associations. The propensity score-matched analysis of initiators compared with non-initiators resulted in a point estimate similar to the crude estimate but with a large CI [HR 0.50 (0.27, 0.93)].

Estimates of associations between statin use and non-CVD mortality were comparable between early users (1990–95) and late users (1996–98) for methods that compared prevalent users to non-users (Table 2). Methods that compared initiators to non-initiators, including the propensity score-matched model, produced more protective associations among early initiators than late initiators.

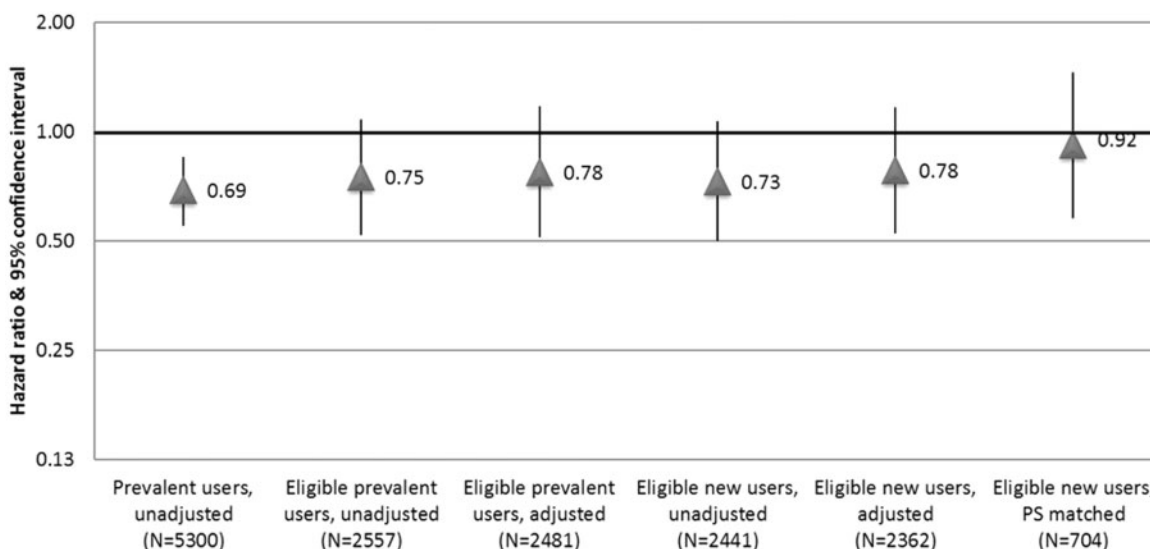
We found no evidence of effect modification by age, sex or propensity of treatment. Results were similar when we examined the patterns for risk differences (Supplementary Table 2, available as Supplementary Data at *IJE* online). Estimates from the inverse probability of treatment weighted model were similar to those from the propensity score-matched model. Adjustment for inverse probability of attrition weights did not have an impact on the estimates.

**Table 1.** Characteristics of participants in the Cardiovascular Health Study by statin status

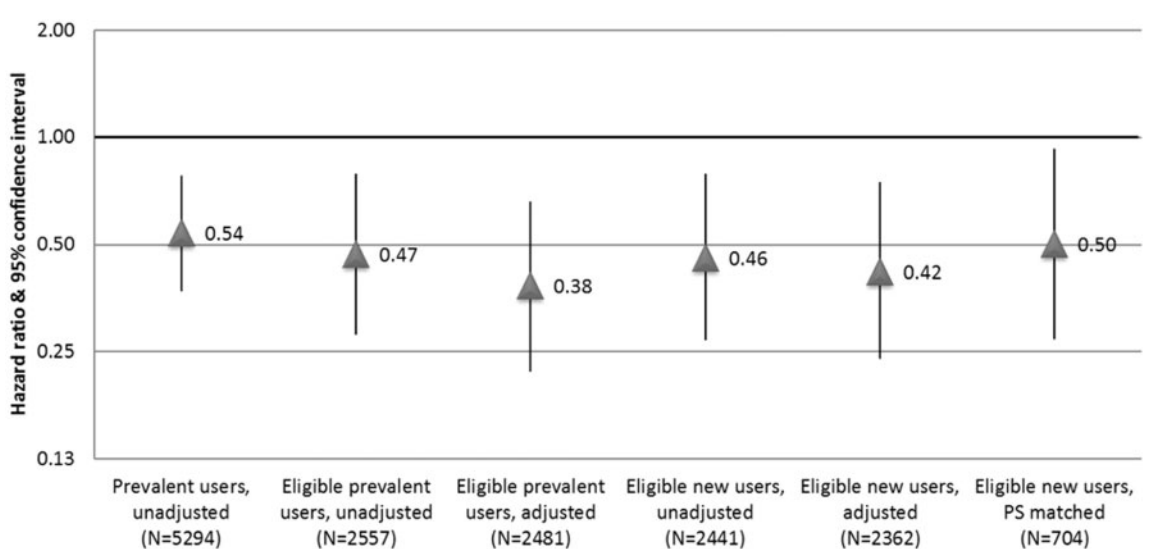
	Non-users	Users	Initiators
Number of observations	39 762	2493	697
Number of individuals	5282	727	628
<b>From all pooled observations:</b>			
Age	76.2	75.9	75.9
Female	60.9%	69.8%	66.7%
Non-White	13.5%	12.6%	12.8%
LDL-cholesterol	127.7	141.5	149.5
BMI	26.7	27.2	27.3
Current smoker	9.3%	8.0%	8.2%
Prevalent CHD	15.4%	35.6%	37.0%
Diabetic	14.5%	16.3%	16.8%
Taking OHGA	6.6%	9.9%	10.3%
Private insurance	65.0%	63.7%	65.4%
4+ years of college	21.7%	23.1%	20.6%
Income > \$35 000	24.7%	32.6%	29.1%
Eligible for statins	32.3%	59.7%	55.7%
% of observations with incident MI within 5 years	6.6%	3.7%	4.3%
% of observations with non-CVD death within 5 years	12.6%	9.0%	10.1%

## Discussion

Using a well-characterized cohort of older adults followed in the years after the introduction of statins, we found that a variety of design and analytic methods to address bias in observational data had a modest effect on the results. For non-cardiovascular mortality, matching on exposure propensity scores in an eligible new-user population resulted in a point estimate that approximated the findings from RCTs. In contrast, for the outcome of MI, all estimates were further in the protective direction compared with the observed effect sizes in RCTs. Despite a sample size of more than 5000 participants, sample size appeared to be a limiting factor due to the poor precision of the estimates. Additionally, though we observed a large influx of older new users, early adopter bias appeared to be a persistent limitation, especially for the outcome of MI.



**Figure 1.** Comparing approaches to estimate the association between statin use and non-cardiovascular mortality in the Cardiovascular Health Study.



**Figure 2.** Comparing approaches to estimate the association between statin use and incident MI in the Cardiovascular Health Study.

**Table 2.** HR (95% CI) between statin use and non-cardiovascular mortality among early (1990–95) vs late (1996–98) study visits

Model	1990–95	1996–98
Prevalent users, unadjusted	0.60 (0.41, 0.87)	0.58 (0.46, 0.73)
Eligible prevalent users, unadjusted	0.70 (0.38, 1.30)	0.56 (0.37, 0.83)
Eligible prevalent users, adjusted	0.75 (0.39, 1.44)	0.79 (0.50, 1.26)
Eligible new users, unadjusted	0.44 (0.22, 0.88)	0.81 (0.52, 1.27)
Eligible new users, adjusted	0.49 (0.24, 1.00)	1.07 (0.64, 1.79)
Eligible new users, PS-matched	0.37 (0.15, 0.94)	1.25 (0.73, 2.14)

Restriction to participants eligible for statins, though important in other settings,<sup>6–9,22</sup> did not have a distinguishable impact on the estimates from our study. Similarly, comparing new users to eligible non-initiators and multivariable adjustment did not result in substantial changes to point estimates. There were few prevalent statin users in CHS, particularly at the beginning of the study, minimizing the difference between analysis of all statin users and analysis of statin initiators. Previous research has suggested that covariate adjustment may be less effective than restriction for addressing healthy user bias<sup>9,33</sup>; the lack of effectiveness in this setting may reflect the persistence of early adopter bias present.

All estimates of the association between statin use and incident MI were more protective than those from a randomized trial in a similarly aged (72–80 years) population [PROSPER, HR 0.81 (95% CI: 0.69, 0.94) for CHD death or non-fatal MI],<sup>25</sup> a general meta-analysis [RR 0.73 (0.67, 0.80) for fatal and non-fatal CHD events]<sup>17</sup> and a recent subgroup analysis of the ALLHAT-LTT trial aged 65 and older [HR 0.80 (0.62, 1.04) for CHD events].<sup>34</sup> Based on differences in participant characteristics and adherence in CHS and PROSPER, we would expect statins to be less effective in CHS than in PROSPER: adherence to statins was lower in CHS than in PROSPER and CHS participants were healthier than PROSPER participants. The strongly protective associations observed in CHS likely reflect healthy user bias, which was not ameliorated by the design or analytic strategies used here. In particular, we found evidence of a more protective association between statins and non-CVD mortality among those who initiated statins between 1990 and 1995 compared with those who initiated later. This may suggest the presence of early adopter bias—an exaggerated form of healthy user bias that may be particularly difficult to address.

The design and analytic strategies used in this analysis appeared to be more effective in reducing bias in the association between statins and non-CVD mortality than in the association between statins and incident MI. This differential effectiveness may reflect a nuance of healthy user bias, such that CHS participants who used statins also took other cardio-protective measures to preserve their cardiovascular health, but that did not improve their non-CVD risk. In other words, the ignorability assumption may be met for one outcome, but not another. This may also explain, as others have noted, why confounding may be less important for studies of adverse effects in which the risk factors for adverse effects are different than the reasons for treatment.<sup>16</sup> We did not observe an effect of adjustment for loss to follow-up via inverse probability of attrition weights, suggesting that either (i) attrition did not have important effect on the estimates of association or (ii) attrition

could not be accurately modelled given the observed covariates.

The success of similar design and analytic approaches in replicating RCT results on effectiveness outcomes<sup>5–7</sup> may underscore the importance of early adopter bias. CHS's timing, with recruitment beginning shortly after the release of statins to the market, highlights early adopter bias as a particularly pervasive form of healthy user bias, for which these methods may be insufficient. Another possibility is that CHS's smaller sample size than the datasets used in previous analyses prevented us from observing the true impact of these methods. However, small sample size would not account for the reduced bias in associations with non-CVD mortality but not with incident MI. Much additional research is needed to evaluate these methods in a variety of settings and consider the implications for using observational studies immediately after approval of new therapies as post-marketing surveillance.

Some limitations affect the interpretation of these results. Our ability to measure eligibility for statins was limited by the infrequently-measured LDL-cholesterol and irregular timing of visits in relation to actual initiation of statins, which may partially explain why restriction to those eligible for statins did not affect point estimates substantially. Only 60% of statin initiators were classified as eligible based on covariate values at the visit prior to initiation, likely reflecting both measurement error and that prescribing practices were not strictly in line with contemporary guidelines. Additionally, we did not include details of statin type or dosage, which could affect the expected magnitude of benefit, as we did not have enough observations to compare across subgroups. We used an untreated comparison group; active comparator groups may be more useful for addressing bias, though finding an appropriate comparator to address early adopter bias in statin users would be challenging. Finally, adherence in CHS was worse than in PROSPER (as in most observational studies, compared with randomized trials), which would lead to bias towards the null. We did not have information on whether participants started (or stopped) taking statins between yearly visits, likely increasing measurement error, though this would also lead to bias towards the null. We operationalized statin initiation as precisely as possible, though we do not capture those who initiated and quit statins between study visits. This may have led to an overestimation of adherence, especially among those who experienced a statin-associated adverse event. The strengths of this analysis include the use of a well-characterized, population-based cohort to observe uptake of a novel therapy. By using two outcomes of interest, including a negative control outcome, and leveraging results from a relevant randomized trial, we were able to

draw conclusions about the capacity for various methods to address bias in observational data.

This analysis illustrates the complications of drawing causal inferences from observational data. Although the combination of design and analysis produced expected estimates of the effect of statins on non-CVD mortality, the magnitude of association between statins and incident MI remained further in the protective direction than expected, suggesting residual early adopter bias. Additionally, poor precision of the estimates remained a limitation due to limited events among statin users, despite a sample size of more than 5000 participants and 5-year risks of 31% for non-CVD death and 14% for MI. Future real-world examples that evaluate the conditions under which estimates from observational data can approximate trial data will help elucidate the most common strengths and limitations of these methods, and better inform meaningful application of these methods.

## Supplementary Data

Supplementary data are available at *IJE* online.

## Funding

This research was supported by the National Institute on Aging (K01AG047273, K01AG039387, R01AG046206). Support for CHS was provided by contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295 and U01HL130114 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Conflict of interest:** B.P. serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the steering committee of the Yale Open Data Access Project.

## References

- Rosenbaum PR. From association to causation in observational studies: the role of tests of strongly ignorable treatment assignment. *J Am Stat Asso* 1984;79:41–48.
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
- Pearl J. Causal inference in statistics: an overview. *Statistics Surveys* 2009;3:96–146.
- Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med* 2007;26:20–36.
- Hernán MA, Alonso A, Logan R *et al*. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology (Cambridge, Mass)* 2008;19:766.
- Danaei G, Rodríguez LAG, Canero OF, Logan R, Hernán MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res* 2013;22:70–96.
- Schneeweiss S, Patrick AR, Stürmer T *et al*. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. *Med Care* 2007;45(Suppl 10):S131.
- Psaty BM, Siscovick DS. Minimizing bias due to confounding by indication in comparative effectiveness research: the importance of restriction. *JAMA* 2010;304:897–98.
- McGrath LJ, Ellis AR, Brookhart MA. Controlling time-dependent confounding by health status and frailty: restriction versus statistical adjustment. *Am J Epidemiol* 2015;kwu485.
- Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003;158:915–20.
- Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res* 2011;46:399–424.
- Oakes JM, Johnson PJ. Propensity score matching for social epidemiology. *Methods in Social Epidemiology* 2006;1:370–93.
- Silverman SL. From randomized controlled trials to observational studies. *Am J Med* 2009;122:114–20.
- Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. *Am J Prev Med* 2007;33:155–61.
- Miettinen OS. The need for randomization in the study of intended effects. *Stat Med* 1983;2:267–71.
- Vandenbroucke JP. When are observational studies as credible as randomised trials? *The Lancet* 2004;363:1728–31.
- Taylor F, Huffman MD, Macedo AF *et al*. Statins for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev* 2013;1:CD004816. doi: 10.1002/14651858.CD004816.pub5.
- Desai CS, Martin SS, Blumenthal RS. Non-cardiovascular effects associated with statins. *BMJ* 2014;349:g3743.
- Mihaylova B, Emberson J, Blackwell L *et al*. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet* 2012;380:581.
- Majumdar SR, McAlister FA, Eurich DT, Padwal RS, Marrie TJ. Statins and outcomes in patients admitted to hospital with community acquired pneumonia: population based prospective cohort study. *BMJ* 2006;333:999.
- Dormuth CR, Patrick AR, Shrank WH *et al*. Statin adherence and risk of accidents a cautionary tale. *Circulation* 2009;119:2051–57.
- van Rein N, Cannegieter SC, le Cessie S, *et al*. Statins and risk of bleeding: an analysis to evaluate possible bias due to prevalent users and healthy user aspects. *Am J Epidemiol* 2016;kwv255.
- Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. *Pharmacoepidemiology and Drug Safety* 2010;19:858–68.



24. Lemaitre RN, Psaty BM, Heckbert SR, Kronmal RA, Newman AB, Burke GL. Therapy with hydroxymethylglutaryl coenzyme a reductase inhibitors (statins) and associated risk of incident cardiovascular events in older adults: evidence from the Cardiovascular Health Study. *Arch Intern Med* 2002;**162**: 1395–1400.
25. Shepherd J, Blauw GJ, Murphy MB, *et al.* Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *The Lancet* 2002;**360**:1623–30.
26. Fried LP, Borhani NO, Enright P *et al.* The cardiovascular health study: design and rationale. *Ann Epidemiol* 1991;**1**:263–76.
27. Carroll RJ, Day N, DeMets DL *et al.* Discussion on “Statistical issues arising in the women’s health initiative”. *Biometrics* 2005; **61**:911–35.
28. Zeger SL, Liang K-Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;**42**:121–30.
29. Grundy SM, Bilheimer D, Chait A *et al.* Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel II). *JAMA* 1993;**269**:3015–23.
30. Psaty BM, Lee M, Savage PJ *et al.* Assessing the use of medications in the elderly: methods and initial experience in the Cardiovascular Health Study. *J Clin Epidemiol* 1992;**45**:683–92.
31. Ives DG, Fitzpatrick AL, Bild DE *et al.* Surveillance and ascertainment of cardiovascular events: the Cardiovascular Health Study. *Ann Epidemiol* 1995;**5**:278–85.
32. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat* 1985;**39**:33–38.
33. Lalmohamed A, van Staa TP, Vestergaard P *et al.* Statins and risk of lower limb revision surgery: the influence of differences in study design using electronic health records from the United Kingdom and Denmark. *Am J Epidemiol* 2016;**kwv311**.
34. Han BH, Sutin D, Williamson JD *et al.* Effect of statin treatment vs usual care on primary cardiovascular prevention among older adults: the ALLHAT-LLT randomized clinical trial. *JAMA Internal Medicine* 2017;**177**:955–65.