

# The Active Comparator, New User Study Design in Pharmacoepidemiology: Historical Foundations and Contemporary Application

Jennifer L. Lund<sup>1</sup> · David B. Richardson<sup>1</sup> · Til Stürmer<sup>1</sup>

Published online: 30 September 2015  
© Springer International Publishing AG 2015

**Abstract** Better understanding of biases related to selective prescribing of, and adherence to, preventive treatments has led to improvements in the design and analysis of pharmacoepidemiologic studies. One influential development has been the “active comparator, new user” study design, which seeks to emulate the design of a head-to-head randomized controlled trial. In this review, we first discuss biases that may affect pharmacoepidemiologic studies and describe their direction and magnitude in a variety of settings. We then present the historical foundations of the active comparator, new user study design and explain how this design conceptually mitigates biases leading to a paradigm shift in pharmacoepidemiology. We offer practical guidance on the implementation of the study design using administrative databases. Finally, we provide an empirical example in which the active comparator, new user study design addresses biases that have previously impeded pharmacoepidemiologic studies.

**Keywords** Pharmacoepidemiology · Confounding · Selection bias · Study design

---

This article is part of the Topical Collection on *Pharmacoepidemiology*

✉ Jennifer L. Lund  
Jennifer.Lund@unc.edu

David B. Richardson  
david.richardson@unc.edu

Til Stürmer  
sturmer@unc.edu

<sup>1</sup> Department of Epidemiology, University of North Carolina at Chapel Hill, McGavran-Greenberg Hall, 2102-D, CB #7435, Chapel Hill, NC 27599-7435, USA

## Introduction

Over the past 40 years, pharmacoepidemiologists have made substantial contributions to public health by generating evidence regarding the use, benefit, and harm of drugs, biologics, vaccines, and medical devices in the population. During this time, there have been rapid methodological advances within the field. In this paper, we review biases that can affect pharmacoepidemiologic research and one of the most influential methodological developments used to address these biases: the active comparator, new user (ACNU) study design. We review the study design’s historical foundations, provide practical guidance for its implementation using administrative databases, and examine the study design’s benefits and limitations in practice using a contemporary example. For brevity throughout the remainder of the review, we will refer to any drug, biologic, vaccine, medical device, or any other medical intervention simply as the “treatment.”

## Potential for Bias in Pharmacoepidemiology

### Confounding

In routine clinical practice, treatments are administered to patients for a specific reason (i.e., the medical indication)—to prevent an adverse health outcome from occurring or to treat an adverse health outcome that has already occurred (and to prevent future occurrences). The indication for treatment leads to selective prescribing of treatments, and in pharmacoepidemiologic studies, confounding by indication may arise if the indication for the treatment is also related to prognosis (the risk of the outcome) [1, 2]. Confounding by indication can be illustrated using a hypothetical study seeking to estimate the effect of beta-agonists on asthma-related

mortality among asthma patients. Asthma patients with severe disease are more likely than patients with less severe disease to receive beta-agonists and to die from their asthma. Such confounding would tend to make beta-agonists appear as though they were associated with asthma mortality.

In general, confounding by indication is likely to be of greatest concern in studies that compare initiators of treatments to non-initiators. Schneeweiss and colleagues provide support for this concern [3] in a study evaluating the association between statin use and 1-year mortality among adults aged 65+ years. Through incremental restriction of the study cohort, the authors demonstrate a large shift in the overall estimate when new statin users were compared with new users of an unrelated preventive treatment (anti-glaucoma medication) instead of non-users, indicating a large reduction in bias.

Another type of confounding bias common in pharmacoepidemiologic studies is the healthy initiator bias [4–6]. This bias can arise through two distinct pathways [7, 8]. The first involves the selective initiation of preventive treatments among healthy and health-conscious patients, who, through the effects of their healthy lifestyle, are also at decreased risk of a number of adverse health outcomes. The second involves the selective channeling of treatments away from frail individuals, who are at an increased risk of adverse outcomes. Under both scenarios, the healthy initiator bias will lead to spurious associations, where the beneficial effect of a given drug will be exaggerated.

The healthy initiator bias has been documented in studies of influenza vaccine effectiveness in older adults and other populations in poor health [9, 10]. Older adults and particularly those with severe comorbid illness and decreased physiologic reserves are unlikely to receive influenza vaccination but are at an increased risk for death. This type of channeling has led a number of non-experimental studies to report implausibly strong preventive effects of influenza vaccination on all-cause mortality in the range of a 50 % relative risk reduction [11–13].

### Selection Bias

One common form of selection bias that occurs in pharmacoepidemiologic studies is the healthy adherer bias [6, 14, 15, 16, 17], which extends the healthy initiator and frailty bias to patients who adhere to treatment. Individuals who adhere to preventive treatments over prolonged periods are as follows: (1) more likely to adhere to other healthy behaviors and preventative care and (2) less likely to have experienced changes in frailty (or underlying health status) compared with their non-adherent counterparts. One of the most striking examples of the healthy adherer bias was demonstrated in a meta-analysis evaluating the effect of adherence to

placebo in randomized controlled trials (RCTs) [18]. This study reported that patients who adhered to placebo had substantially reduced mortality compared with patients who were non-adherent to placebo. As placebo could not plausibly have any effect on mortality, the observed reduction in mortality can be solely attributable to selection bias.

The impact of (1) confounding by indication and (2) the healthy user bias (the combination of the healthy initiator and healthy adherer biases) works in opposite directions. Confounding by indication distorts drug-outcome associations so that treatment looks “bad” or “harmful,” while the healthy user bias distorts drug-outcome associations so that treatment looks “good” or “beneficial.” The relative importance of these biases depends upon the specific drug-outcome association and population of interest.

### A New Standard: the Active Comparator, New User Design

The ACNU design has been one of the most influential methodological advances in pharmacoepidemiology. This study design was developed to avoid many of the biases mentioned above and is regarded as the standard for pharmacoepidemiology, to be implemented whenever possible [19]. In brief, the ACNU design emulates the intervention part of a RCT. Cohorts of new drug users (i.e., individuals newly prescribed an index drug A and individuals newly prescribed a therapeutic alternative or comparator drug B) are assembled and followed over time for the health outcome(s) of interest (Fig. 1, top panel). Notably, this design does not compare treatment users or initiators to non-users.

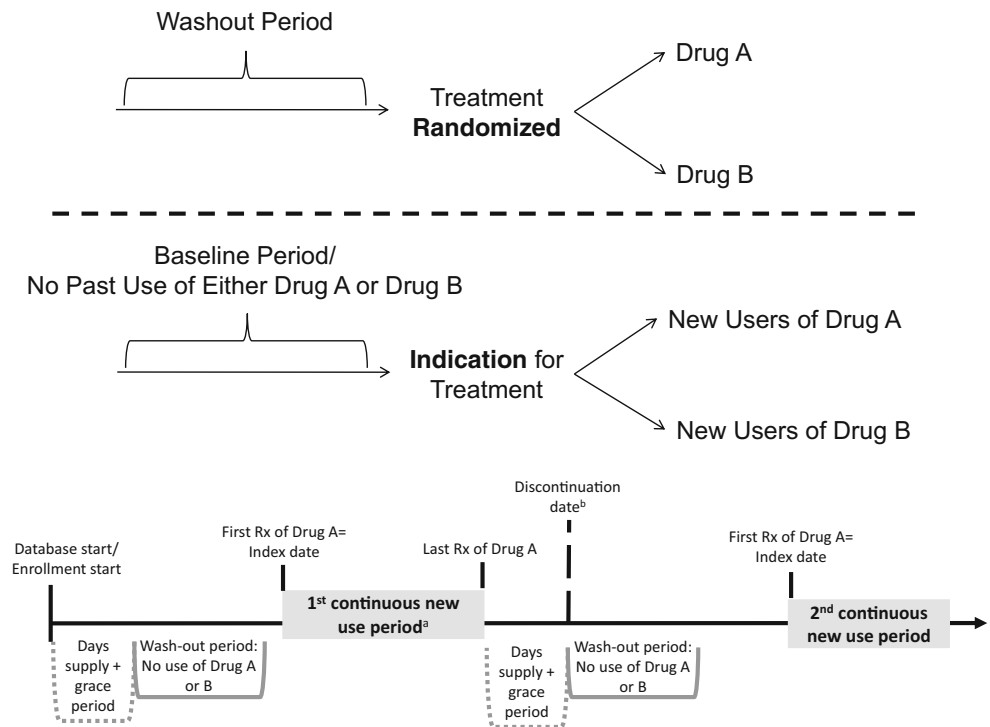
The active comparator (AC) component of the design helps to mitigate bias by restricting the study to individuals with an indication for treatment and without contraindications, including frailty, while the new user (NU) component mitigates bias by aligning individuals at a uniform point in time to start follow-up (i.e., treatment initiation) and ensuring the correct temporality between covariate and exposure assessment. The foundations of this study design have evolved over time, and we first review a number of the contributions from the literature.

### Historical Foundations of the Active Comparator, New User Study Design

#### Contributions from Occupational Epidemiology

In occupational epidemiology, the healthy worker bias [20–22] arises through similar mechanisms as the healthy user

**Fig. 1** Active comparator, new user (ACNU) study design schematics. The *top panel* illustrates how the ACNU study is designed to emulate a head-to-head randomized controlled trial. The *bottom panel* provides a detailed picture of how to identify periods of new use of drug A (the same process would apply to drug B) in a claims or other health care database. <sup>a</sup>One individual can have multiple new use periods. The individual can also be a new user of drug A and later a new user of drug B (or vice versa). Often, analyses will be restricted to the first period of new use. <sup>b</sup>The date of discontinuation (or switching or augmenting) may be used as a censoring date in as-treated analyses



bias, whereby healthy individuals are more likely to become employed (healthy hire bias) [23, 24] and remain in the workforce (healthy worker survivor bias) [20, 21] compared with less healthy individuals. Healthy hire bias is typically controlled by restricting the study to individuals who are employed, comparing disease occurrence between subgroups of workers within the study [21, 25], lending early support to the use of ACs in epidemiologic studies.

The healthy worker survivor bias is a form of selection bias when changes in underlying health status cause people to leave the workforce [21, 26]. Those who remain employed over prolonged periods of time are therefore an increasingly selected group of people who stayed healthy rather than a random sample of all those who entered the workforce. This selection bias can be minimized by starting follow-up at hire (referred to as an inception cohort) and not conditioning on prolonged employment or addressed analytically, assuming that information on time-varying predictors of continued employment is available.

**Contributions from Clinical Epidemiology and Pharmacoepidemiology**

*Active Comparator*

In a paper published in 1987, Kramer and colleagues [27••] highlighted a number of methodological issues affecting non-

experimental studies of adverse drug effects. The authors succinctly recommended the use of an AC to avoid confounding in pharmacoepidemiologic studies, stating that “Even if it is clear to whom a particular drug poses a risk, and over what period of time, it is important to compare that risk with that of some other real therapeutic option for patients with the same clinical indication.” Kramer recommends that “... any epidemiologic study of treatment risks should compare two or more viable treatment alternatives,” and “... measuring risks conditionally on clinical indication is...essential to reduce confounding.”

In a 1989 paper [28], Ray and Griffin further the suggestion to “...select controls [i.e., comparators] from persons receiving similar drugs. In a case-control study, this can be achieved by comparing the odds ratio for these other drugs to that of the drugs under consideration.”

Gerstman and colleagues also support the use of ACs to control for confounding by indication in a 1990 paper [29] stating, “One means of approximating a match based on indication is to restrict the study base to individuals using drugs within a single therapeutic category. Rates of suspected adverse reactions may then be compared among alternative therapies presumably used to treat similar underlying conditions.” The authors continue that “Studies restricted to patients using drugs within a particular class may also aid in controlling for selection and detection biases.”

## New User

One of the central tenants of a NU design is that the start of follow-up should be anchored to the initiation of a treatment. This concept of a “time zero,” the date of treatment initiation, was first raised by Feinstein in 1971 [30] in writings on “chronology bias,” a general discussion of biases impacting studies of disease prognosis and medical interventions. Feinstein advocates for the use of “inception cohorts” because the use of survivor cohorts “create major distortions in cohort statistics unless the investigator ascertains that the [actual] inception cohorts would have been similar in their short-term (and intermediate) survival rates.”

Kramer and colleagues further the rationale for a NU approach [27•], arguing that “The risk posed by a drug for a particular event is not generally the same in the sixth month of chronic therapy as in the first or second week. Using the total time exposed as the denominator in estimating risk is therefore incorrect, particularly if many patients are on chronic therapy... The two approaches are equivalent only if the risk remains constant as a function of time, i.e., the timing distribution is uniform and the expected time per course of therapy is known. The first of these criteria is almost never true, and the second is clearly untrue for the drugs under study...”

In 1989, Guess expanded on the issue of time-varying hazards for drug side effects after drug initiation [31]. Guess states that “The possibility of temporally non-constant hazard functions should be considered in the study design. *This requires that drug exposure time be measured not only in relation to onset of the study disease but also in relation to start of therapy with the study drug [italics by author].*” Guess continues that “If the duration-specific incidence ratios differ widely, the odds ratio will depend not only on the drug but also on its usage pattern in the study population.”

In 1994, Moride and Abenheim [32] highlight the concept of depletion of susceptibles (i.e., selection bias, to explain time-varying hazards, also mentioned by Guess [31]). The authors concluded that their counterintuitive results indicated “...either a lack of cumulative effect of the drugs, or more likely, a selection process by which patients who have used the drugs in the past and tolerated them well remain on the drugs while patients who are susceptible to gastropathy select themselves out of the population at risk. This process is analogous to the well-known “healthy worker effect”.... If not taken into account in comparative studies, it could introduce a selection bias.”

Throughout the 1990s, additional research underscored concerns regarding potential for bias due to time-varying hazards and the mixing of new and prevalent users with

contributions by Gerstman et al. [29], Hartzema [33], and McMahon and colleagues [34, 35]. In the most cited paper (479 citations, according to Web of Science (August 19, 2015) on the topic published in 2003 in the American Journal of Epidemiology, Ray coined the phrase “new user design” [36•]. In addition to reiterating the above pitfalls of time-varying hazards, Ray mentions, for the first time, a concern regarding the timing of covariate assessment where “...covariates for [prevalent] drug users at study entry often are plausibly affected by the [earlier] drug itself.” He then argues that the “new-user design eliminates these biases by restricting the analysis to persons under observation at the start of the current course of treatment.” Ray’s landmark paper focuses on the theory and application of the NU design in the cohort setting and has led to wide dissemination of the underlying message and substantial changes in best practices in pharmacoepidemiology [19].

## Practical Guidance on Implementation of the Active Comparator, New User Design

When planning a study using the ACNU design, it is necessary to carefully consider the following: (1) the selection of an AC, (2) the definitions used to identify “new use,” and (3) treatment changes over time. We review each of these considerations below and provide practical advice and an illustration to aid implementation for pharmacoepidemiologic research.

### Selection of an Active Comparator

The purpose of selecting an AC is to mitigate confounding by indication and other unmeasured patient characteristics (e.g., healthy initiator, frailty). Therefore, the more “substitutable” one treatment is for another, the less potential unmeasured confounding. Clinical treatment guidelines are an obvious resource to consult when selecting an AC for an index treatment of interest. Another option is to evaluate the similarity of measured characteristics of patients initiating the treatment of interest and each of the potential AC treatments and select the comparator that most closely matches the treatment of interest with respect to, e.g., age, sex, comorbidity, and concomitant medication use. If none of the measured factors is a strong predictor of initiating the treatment of interest rather than the comparator, it is plausible that unmeasured factors would also not affect treatment choice.

### Identifying New Users

NUs of a treatment have initiated therapy after some defined period of non-use and are at the beginning of their treatment course. Prevalent users include both NUs and current users of a treatment and are observed at a number of different time

points in their treatment course (i.e., there is no uniform time zero). By including only NUs in pharmacoepidemiologic studies and starting follow-up at treatment initiation, time-varying hazards can be described and evaluated and the temporality of covariate assessment is preserved.

Ideally, NUs would be first-time-ever users of a treatment. To assess this type of new use, lifetime treatment data would be necessary (e.g., Danish prescription databases). In most settings, however, defining new use requires consideration of a “washout period” (Fig. 1, top panel). This period represents a fixed window prior to the first prescription (i.e., treatment initiation) where the individual has drug coverage but does not have any use of either the index treatment or the AC treatment. Washout windows of 6–12 months are common, although evaluation of longer periods has revealed misclassification of “true” new use with relatively short washout periods [37].

Lack of lifetime treatment data when implementing an ACNU design results in the possibility for individuals to have multiple new use periods. This would occur when an individual initiates treatment, but subsequently discontinues (i.e., stops refilling the prescription) and re-initiates after the washout period. Because the exact date of treatment discontinuation is generally unknown, the discontinuation date is usually defined as the last prescription dispensing date + the days supply + a grace period. A person refilling a prescription of the same drug before the discontinuation date is classified as a continuous user. The grace period is necessary to allow for some leeway when refilling prescriptions, for patients forgetting to take some pills, or even splitting pills. It is frequently based on a proportion (e.g., 50 %) of the typical day’s supply of the drug but can also be based on the empirical distribution of days between refills. Researchers interested in allowing for multiple new use periods should add the length of the day’s supply plus the grace period to the initial washout window in order to provide uniform washout windows for all new use periods (Fig. 1, lower panel). They should also index the new use periods per patient so that the first one can easily be identified if an analysis restricted to the first new use period is desired (first new use period could be the first-ever use period for some individuals in contrast to subsequent ones). Once new use is established, the index date is defined as the first prescription dispensing date after being free of both the index and comparator treatments during the washout period. In some settings, a second prescription within a given amount of time from the first prescription is required for cohort entry, in which case, the index date would be the second prescription dispensing date.

By requiring the start of follow-up to occur at the point of treatment initiation, analysis of the changing hazards of events of interest is straightforward. For example, the rate of an adverse event can be evaluated among the index and AC cohorts in the first 30, 60, or 90 days following therapy initiation. If

the event hazard is highest in the early periods and decreases in later periods, this may be an indication of the depletion of susceptibles.

### Analysis of Treatment Changes Over Time

An additional advantage of the ACNU design is that researchers can avoid the potential for the healthy adherer bias by ignoring treatment changes over time (e.g., if stopping is ignored, then no selection bias induced.) This “first treatment carried forward” approach is the equivalent of the “intent to treat” analysis in RCTs. While this approach avoids selection bias, another bias, exposure misclassification, is introduced and would tend to increase over time since treatment initiation. This analysis is usually seen as introducing bias toward the null and is often acceptable in conventional RCTs because it is conservative when assessing efficacy. When assessing adverse outcomes, however, as in many pharmacoepidemiologic studies, bias toward the null may mask potential harm, is therefore not conservative, and should be avoided. It is also worth mentioning that in RCTs where major crossover between treatments is not assumed, bias is very likely to be toward the null. With major crossover between treatments, it would be possible for bias to cause the treatment effect estimate to cross the null.

An alternative to the first treatment carried forward design is the “as-treated” analysis, the parallel to the “per-protocol” analysis in RCTs, where we censor patients at the time of treatment changes. This analytic approach in pharmacoepidemiologic studies can account for not only stopping treatments (described above in Fig. 1, lower panel), but also switching of treatments (switching from drug A to drug B) and adding on treatments (adding drug B to drug A and vice versa). The distinction between switching and augmenting can only be made after a period of time, because it will depend on whether or not the individual refills the initial treatment after start of the alternative treatment. The period of time that should be used to assess a refill of the initial treatment should be based on, e.g., the day’s supply of the previous prescription and the grace period.

Once these events (starting, stopping, switching, augmenting) have been identified for each patient, researchers must then hypothesize about three distinct periods to define time-at-risk periods: the induction period (i.e., time required for the treatment to biologically effect the outcome), latent period (i.e., time required between outcome onset and clinical diagnosis), and potential carryover effects (i.e., time that the treatment remains active in the body) based on available information on the treatment and outcome of interest. In practice, carryover effects tend to be short, only accounting for a few days, and therefore are generally ignored or assumed to be part of the latent period following discontinuation.



These concepts can be illustrated using two hypothetical examples. First, consider a study of the effect of an antibiotic on the adverse outcome of anaphylaxis. In this case, we can assume that treatment has an immediate effect on the incidence of anaphylaxis (i.e., an induction period=0 days) and is instantly diagnosed by a health care provider (i.e., latent period=0 days), and as such, we would start time-at-risk at drug initiation. Furthermore, if we assume that the drug is eliminated from the body immediately upon stopping (i.e., carryover period=0 days), then time-at-risk would stop at drug discontinuation. Note that we would likely miss any effect if we did not implement a NU design.

Second, consider a study of the effect of antidepressant medications on risk of cancer. We may hypothesize that antidepressants have a late-acting effect in the carcinogenesis process (e.g., relatively short induction period, say 6 months) and would need to add a latent period for cancer to be clinically detected (e.g., an additional 6 months). It is likely that the carryover effect of antidepressants is relative short (only a few days). As such, we may consider a combined induction and latent period of 12 months and stop follow-up 6 months after drug use has ended (i.e., assigning person-time and events up to 6 months after the actual stopping date). Sensitivity analysis varying the length of the assumed induction, latent, and carryover periods, ideally in both directions (longer and shorter), can help assess robustness of estimated effects and is recommended.

Finally, the ACNU design may aid in the evaluation of the cumulative effects of drug exposures on health outcomes. Because all individuals in the cohort are followed from drug initiation, the event rate in those who continue on therapy on drug A (e.g., for 2 years) can be compared to individuals who continue therapy on drug B (for 2 years). Consideration of the mechanisms and patterns of treatment stopping, switching, and augmenting is important for the validity of this approach, as informative censoring, where individuals who are censored from the analysis are at increased (or decreased) risk of the outcome of interest, may still induce selection bias, especially if differential between treatment cohorts. Analytic approaches that account for informative censoring may be applied, if detailed information is available on the predictors of non-adherence. Many of these methods have been developed and gained popularity in both the occupational epidemiology literature [26] and the HIV literature [38, 39], perhaps based on the ability to predict reasons for employment cessation and changes in antiretroviral treatments, respectively.

### Benefits of the Active Comparator, New User Design: a Contemporary Example

To illustrate the benefits of the ACNU design, we will briefly review a study by Stürmer and colleagues evaluating the long-

term effects of insulin glargine on cancer risk [40]. One of the major challenges in evaluating the effects of insulin on cancer risk in patients with type 2 diabetes is likely strong confounding by indication (diabetes severity and duration) and obesity, which are impossible or difficult to measure in administrative data.

Using clinical guidelines, the authors selected the second most prescribed long-acting insulin, human NPH insulin, as the AC. The researchers defined a washout period of 19 months to identify NUs of the study drugs, to represent a typical 30-day supply of insulin + a grace period of 6 months + a washout period of 12 months, using an approach similar to Fig. 1. The relatively long washout period was derived from the empirical distribution of days between refills, indicating that any shorter period would have led to misclassify many apparently continuous users as having stopped treatment.

Potential confounders were assessed in the 12 months prior to the first prescription for insulin, which ensures the correct temporality of covariate assessment in relation to the drug initiation (i.e., prior to initiation). The primary analysis used an as-treated approach where follow-up began at the date of the second prescription and stopped after treatment discontinuation, date of treatment switch (i.e., filling a prescription for another long-acting insulin), death, end of enrollment, end of the study period, or the date of a cancer event. Sensitivity analyses varied induction and latent periods (i.e., excluding incident cancer cases for up to 12 months after insulin initiation and allowing for diagnoses to be made up to 24 months or indefinitely after stopping treatment [first treatment carried forward]). Furthermore, the cumulative effects of insulin on cancer risk were assessed by comparing initiators who remained on treatment for 0 to less than 6 months, 6 to less than 12 months, 12 to less than 24 months, and greater than 24 months. The authors found no indication for an increased risk for cancer.

Overall, the measured baseline patient characteristics of initiators of insulin glargine vs. human NPH insulin were similar, with slight differences noted for age, sex, year of treatment initiation, and concomitant medication use. These differences were reduced to a clinically insignificant level or eliminated entirely using propensity score weighting. To address concern about residual confounding by BMI, the authors conducted two validation studies in external populations. In both studies, the distribution of BMI was virtually identical in initiators of insulin glargine and in initiators of human NPH insulin.

The example provides evidence that the ACNU study design can enhance the validity of non-experimental studies through the following: (1) aligning individuals at a uniform point in time to start follow-up (i.e., initiation of a drug), (2) ensuring the correct temporality between covariate and exposure assessment, and (3) mitigating measured and unmeasured confounding by design through the selection of an appropriate

AC that led to the inclusion of patients that had the indication for treatment augmentation.

There are important limitations of this approach worthy of note. “Real-world” treatment dynamics, i.e., few patients stay on the initial insulin over extended periods of time, limits the evaluation of the cumulative effect of treatment beyond 2 years. This is important for interpretation of the study findings and whether concerns over the “longer-term” effects (e.g., 2+years) of treatment exposure on cancer risk are warranted from a public health perspective. Furthermore, data on long-term use available when studying prevalent rather than incident users would not answer a clinically relevant question (we cannot “assign” anyone to long-term use) and would suffer from the selection introduced by focusing on a very small fraction of all those that have ever been exposed to the drug. Another potential limitation of the as-treated approach is the concern about informative censoring. The first treatment-carried-forward analyses that eliminate selection bias support the primary as treated analysis, however. Further work on evaluating the impacts of informative censoring in an ACNU design is indicated. Finally, we need to assume that the BMI balance observed in the external validation studies would be transportable to the main study.

## Conclusions

Advances in our understanding of how non-experimental study design and treatment dynamics impact the validity of pharmacoepidemiologic studies have laid the foundation for the modern implementation of the ACNU design. The ACNU design offers the theoretical advantage of mitigating confounding bias by indication, healthy user, and frailty at the design stage without depending on good measures for these (i.e., prior to and independent of the application of analytic methods). The ACNU design is considered the current standard in pharmacoepidemiologic research that should be implemented whenever possible. The ACNU design can be easily implemented within large automated health care databases. As evidence accrues with respect to the scenarios under which the theoretical benefits are most likely to be actualized, the ACNU design is expected to gain further popularity.

## Compliance with Ethics Guidelines

**Conflict of Interest** Jennifer L. Lund declares that she receives salary support from the UNC Oncology Clinical Translational Research Training Program (5K12CA120780) and research and salary support from the PhRMA Foundation to the Department of Epidemiology at UNC.

David B. Richardson declares no conflict of interest.

Til Stürmer declares research funding from NIA (R01 AG023178, R01 AG042845), the National Cancer Institute (R01 CA174453), and the Patient Centered Outcomes Research Institute (PCORI). He also receives salary support from the Center for Pharmacoepidemiology

(GlaxoSmithKline, UCB BioSciences, Merck) and research support from pharmaceutical companies (Amgen, AstraZeneca) to the Department of Epidemiology, University of North Carolina at Chapel Hill (UNC).

**Human and Animal Rights and Informed Consent** This article does not contain any studies with human or animal subjects performed by any of the authors.

## References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Miettinen OS. The need for randomization in the study of intended effects. *Stat Med.* 1983;2(2):267–71.
2. Csizmadia I, Collet J-P (2006) Bias and confounding in pharmacoepidemiology. In: Strom BL, Kimmel SE editors. *Textbook of pharmacoepidemiology*. West Sussex, England: Wiley, Ltd, doi:10.1002/9781118707999.ch16
3. Schneeweiss S, Patrick AR, Sturmer T, Brookhart MA, Avorn J, Maclure M, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. *Medical care.* 2007;45(10 Supl 2): S131–42.
4. Humphrey LL, Chan BK, Sox HC. Postmenopausal hormone replacement therapy and the primary prevention of cardiovascular disease. *Ann Intern Med.* 2002;137(4):273–84.
5. 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(7576):999.
6. Ray WA, Daugherty JR, Griffin MR. Lipid-lowering agents and the risk of hip fracture in a Medicaid population. *Injury Prevent J Int Soc Child Adolescent Injury Prev.* 2002;8(4):276–9.
7. Brookhart MA, Sturmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. *Med Care.* 2010;48(6 Suppl):S114–20.
8. Sturmer T, Jonsson Funk M, Poole C, Brookhart MA. Nonexperimental comparative effectiveness research using linked healthcare databases. *Epidemiology.* 2011;22(3):298–301. **Epub 2011/04/06.**
9. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol.* 2006;35(2):337–44.
10. Jackson LA, Nelson JC, Benson P, Neuzil KM, Reid RJ, Psaty BM, et al. Functional status is a confounder of the association of influenza vaccine and risk of all cause mortality in seniors. *Int J Epidemiol.* 2006;35(2):345–52.
11. Hak E, Buskens E, van Essen GA, de Bakker DH, Grobbee DE, Tacken MA, et al. Clinical effectiveness of influenza vaccination in persons younger than 65 years with high-risk medical conditions: the PRISMA study. *Arch Intern Med.* 2005;165(3):274–80.
12. Nichol KL, Nordin JD, Nelson DB, Mullooly JP, Hak E. Effectiveness of influenza vaccine in the community-dwelling elderly. *N Engl J Med.* 2007;357(14):1373–81.
13. Nordin J, Mullooly J, Poblete S, Strikas R, Petrucci R, Wei F, et al. Influenza vaccine effectiveness in preventing hospitalizations and deaths in persons 65 years or older in Minnesota, New York, and Oregon: data from 3 health plans. *J Infect Dis.* 2001;184(6):665–70.

14. Brookhart MA, Patrick AR, Dormuth C, Avorn J, Shrank W, Cadarette SM, et al. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. *Am J Epidemiol*. 2007;166(3):348–54. **This is one of the landmark papers documenting the “healthy user bias” in pharmacoepidemiology. Findings from the study indicate that patients who adhere to chronic therapies, such as statins, are also more likely to receive preventive healthcare services, such as cancer screening tests and vaccinations.**
15. Petitti DB. Coronary heart disease and estrogen replacement therapy. Can compliance bias explain the results of observational studies? *Ann Epidemiol*. 1994;4(2):115–8.
16. Rossouw JE. Debate: the potential role of estrogen in the prevention of heart disease in women after menopause. *Curr Control Trials Cardiovasc Med*. 2000;1(3):135–8.
17. Dormuth CR, Patrick AR, Shrank WH, Wright JM, Glynn RJ, Sutherland J, et al. Statin adherence and risk of accidents: a cautionary tale. *Circulation*. 2009;119(15):2051–7. **This paper also contributes compelling evidence of the “healthy user bias” in pharmacoepidemiology. The authors utilize a negative control outcome (i.e., an outcome that cannot plausibly be related to the drug exposure) to demonstrate that patients who adhere to statins are less likely to experience a variety of accidents (e.g., motor vehicle accidents, workplace accidents) than less adherent patients, and as such represent a selected group of more health-seeking individuals.**
18. Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, et al. A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ*. 2006;333(7557):15.
19. Johnson ES, Bartman BA, Briesacher BA, Fleming NS, Gerhard T, Kornegay CJ, et al. The incident user design in comparative effectiveness research. *Effective Health Care Program Research Report No. 32*. (Prepared under Contract No. HHS290200500161). AHRQ Publication No. 11(12)-EHC054-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2012. <http://effectivehealthcare.ahrq.gov/reports/final.cfm>.
20. Checkoway H, Pearce N, Kriebel D. Research methods in occupational epidemiology. *Monographs in epidemiology and biostatistics*, vol. 34. New York: Oxford University Press; 1989.
21. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology*. 1994;5(2):189–96.
22. McMichael AJ. Standardized mortality ratios and the “healthy worker effect”: scratching beneath the surface. *J Occup Med*. 1976;18(3):165–8.
23. Wilcosky T, Wing S. The healthy worker effect. Selection of workers and work forces. *Scand J Work Environ Health*. 1987;13(1):70–2.
24. Flanders WD, Cardenas VM, Austin H. Confounding by time since hire in internal comparisons of cumulative exposure in occupational cohort studies. *Epidemiology*. 1993;4(4):336–41.
25. Choi BC. Definition, sources, magnitude, effect modifiers, and strategies of reduction of the healthy worker effect. *J Occupation Med Offic Pub Industrial Med Assoc*. 1992;34(10):979–88.
26. Buckley JP, Keil AP, McGrath LJ, Edwards JK. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology*. 2015;26(2):204–12.
27. Kramer MS, Lane DA, Hutchinson TA. Analgesic use, blood dyscrasias, and case-control pharmacoepidemiology. A critique of the International Agranulocytosis and Aplastic Anemia Study. *J Chronic Dis*. 1987;40(12):1073–85. **This paper provides a critique of the International Agranulocytosis and Aplastic Anemia Study, a multi-center, population-based case-control study evaluating the association between drug exposures and two adverse outcomes (agranulocytosis and aplastic anemia). This critique is notable because it is one of the first references to a number of important methodological issues in pharmacoepidemiologic studies, including time varying hazards and the use of active comparators.**
28. Ray WA, Griffin MR. Use of Medicaid data for pharmacoepidemiology. *Am J Epidemiol*. 1989;129(4):837–49.
29. Gerstman BB, Lundin FE, Stadel BV, Faich GA. A method of pharmacoepidemiologic analysis that uses computerized Medicaid. *J Clin Epidemiol*. 1990;43(12):1387–93.
30. Feinstein AR. Clinical biostatistics. XI. Sources of ‘chronology bias’ in cohort statistics. *Clin Pharmacol Ther*. 1971;12(5):864–79.
31. Guess HA. Behavior of the exposure odds ratio in a case-control study when the hazard function is not constant over time. *J Clin Epidemiol*. 1989;42(12):1179–84.
32. Moride Y, Abenheim L. Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. *J Clin Epidemiol*. 1994;47(7):731–7.
33. Hartzema AG. Guide to interpreting and evaluating the pharmacoepidemiologic literature. *Ann Pharmacother*. 1992;26(1):96–8.
34. McMahon AD, Evans JM, McGilchrist MM, McDevitt DG, MacDonald TM. Drug exposure risk windows and unexposed comparator groups for cohort studies in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 1998;7(4):275–80.
35. McMahon AD, MacDonald TM. Design issues for drug epidemiology. *Br J Clin Pharmacol*. 2000;50(5):419–25.
36. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915–20. **This paper proposes the new-user design, which eliminates biases related to time-varying hazards and inappropriate assessment of confounders by restricting the analysis to persons who newly initiate treatment. This paper provides clear guidance on the application of the new-user design using administrative claims data and has led to substantial changes in best practices in pharmacoepidemiology.**
37. Riis AH, Johansen MB, Jacobsen JB, Brookhart MA, Stummer T, Stovring H. Short look-back periods in pharmacoepidemiologic studies of new users of antibiotics and asthma medications introduce severe misclassification. *Pharmacoepidemiol Drug Saf*. 2015;24(5):478–85.
38. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656–64.
39. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11(5):561–70.
40. Stummer T, Marquis MA, Zhou H, Meigs JB, Lim S, Blonde L, et al. Cancer incidence among those initiating insulin therapy with glargine versus human NPH insulin. *Diabetes Care*. 2013;36(11):3517–25.