ELSEVIER

Journal of Clinical Epidemiology 66 (2013) S110−S121

Journal of Clinical Epidemiology

Validation sampling can reduce bias in health care database studies: an illustration using influenza vaccination effectiveness

Jennifer Clark Nelson[a,b,*], Tracey Marsh[a,b], Thomas Lumley[b,c], Eric B. Larson[a,d], Lisa A. Jackson[a], Michael L. Jackson[a], and for the Vaccine Safety Datalink Team

[a]*Group Health Research Institute, Seattle, WA 98101, USA*
[b]*Department of Biostatistics, University of Washington, Seattle, WA, USA*
[c]*Department of Statistics, The University of Auckland, Auckland, New Zealand*
[d]*Department of Medicine, University of Washington, Seattle, WA, USA*

Accepted 7 January 2013

Abstract

Objectives: Estimates of treatment effectiveness in epidemiologic studies using large observational health care databases may be biased owing to inaccurate or incomplete information on important confounders. Study methods that collect and incorporate more comprehensive confounder data on a validation cohort may reduce confounding bias.

Study Design and Setting: We applied two such methods, namely imputation and reweighting, to Group Health administrative data (full sample) supplemented by more detailed confounder data from the Adult Changes in Thought study (validation sample). We used influenza vaccination effectiveness (with an unexposed comparator group) as an example and evaluated each method's ability to reduce bias using the control time period before influenza circulation.

Results: Both methods reduced, but did not completely eliminate, the bias compared with traditional effectiveness estimates that do not use the validation sample confounders.

Conclusion: Although these results support the use of validation sampling methods to improve the accuracy of comparative effectiveness findings from health care database studies, they also illustrate that the success of such methods depends on many factors, including the ability to measure important confounders in a representative and large enough validation sample, the comparability of the full sample and validation sample, and the accuracy with which the data can be imputed or reweighted using the additional validation sample information.

Keywords: Aged; Bias (epidemiologic); Comparative effectiveness research; Confounding factors (epidemiology); Influenza vaccines; Propensity score

1. Introduction

Large health care databases are increasingly being used to study treatment effectiveness in medical research [1]. However, using data collected primarily for administrative and clinical purposes to conduct comparative effectiveness research poses many challenges. One major problem is that large databases can have limited ability to characterize important confounding differences in outcome risk between exposed and unexposed persons [2−4]. For instance, database confounder adjustment for health status is often accomplished by broadly defining medical conditions using binary International Classification of Disease (ICD-9) diagnosis codes, or risk score summary measures based on these codes, assigned by the medical provider during patient visits [5−7]. This relatively crude adjustment can lead to residual confounding in effectiveness estimates because ICD-9 codes do not adequately measure disease severity or functional status [4,8−12].

A prominent example of this problem is the estimation of influenza vaccine effectiveness (VE) among the elderly in large database studies, which have consistently found implausibly high-risk reductions against all-cause mortality (~50%) when adjusting only for database information such as binary ICD-9-coded indicators of health status [13−15]. More recent research has indicated that residual confounding may account for some, if not all, of this observed effect [10,11]. Specifically, when examining the association between influenza vaccine and mortality in the preinfluenza control period before the circulation of influenza, even

Conflict of interest: L.A.J. has received research funding from Sanofi Pasteur, Pfizer, Novartis, and GlaxoSmithKline.

* Corresponding author. Biostatistics Unit, Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, USA. Tel.: +206-287-2004; fax: +206-287-2871.

E-mail address: nelson.jl@ghc.org (J.C. Nelson).

0895-4356/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jclinepi.2013.01.015

What is new?

- Use of validation sampling methods, such as imputation or reweighting, can improve the accuracy of comparative effectiveness findings from large health care database studies, which can have limited ability to characterize important confounding differences in outcome risk between exposed and unexposed persons.
- The association between influenza vaccination and risk of all-cause mortality is a useful example for studying confounding in treatment effectiveness studies that rely on administrative databases, as there is strong confounding and a natural control period before influenza season that can be used to assess bias and the ability of more sophisticated methods (like those that use validation sampling) to reduce it.
- The success of validation sampling methods in practice depends on many factors, including the ability to measure important confounders in a large enough validation sample, the comparability of the full sample and validation sample, and the accuracy with which the data can be imputed or reweighted using the additional validation sample information.
- Without clear gold-standard estimates of effectiveness in practice for most exposures, a balance of simulation studies (where truth can be generated) and example applications (where the complexities of real data are present) is needed to more fully understand the optimal implementation and settings for use of validation methods in large health care database studies.

larger reductions in risk (~70%) have been found [10]. Any effect observed during the preinfluenza period represents bias because no association between influenza vaccine and morbidity or mortality is biologically plausible when influenza virus is not circulating. This bias has been shown to be reduced by adjusting for functional limitations obtained from medical chart review [11], which suggests that unmeasured frailty is the most plausible unmeasured confounder in this setting. Such confounding would occur if seniors who are very close to dying are no longer given preventive therapies, such as influenza vaccine.

Although adjusting more comprehensively for additional confounders obtained by medical record review or in-person physical examination has the potential to reduce bias in traditional effectiveness estimates that adjust only for information available in database sources, it may be too expensive to collect these more costly confounders in large database studies, where sample sizes can reach tens or hundreds of thousands. One solution is to collect the more expensive data on a smaller validation sample or a subset of the full database cohort and use validation or two-phase sampling methods to incorporate this information into analyses. Herein, we implement two such approaches, a missing data imputation method and a survey sample reweighting method, to estimate influenza VE in the elderly. We use Group Health Cooperative (GHC) administrative data from a prior influenza VE study [10] (full sample) supplemented by richer confounder data on a subset (validation subsample) that included in-person examinations as part of the Adult Changes in Thought (ACT) study [16]. We use the control time period before influenza season to evaluate each method's ability to successfully reduce confounding bias compared with traditional adjustment approaches that rely solely on confounders from database sources.

2. Materials and methods

2.1. Study design, setting, and population

We used existing cohorts from two prior studies conducted among persons aged 65 years and older who were members of GHC, a managed care organization in Washington State with ~350,000 enrollees. The composition of the GHC population is representative of the surrounding region, which is primarily white, middle class, and well educated. The first was a large, retrospective database cohort study of influenza VE among 72,527 community-dwelling seniors from 1995 to 2002 [10] that captured data from GHC's administrative systems on all-cause mortality (outcome of interest), influenza immunization (exposure of interest), and database confounders used in prior database studies of influenza VE [14,15], including health care utilization (e.g., number of outpatient visits) and ICD-9 diagnosis codes assigned to patient encounters and used to define binary health status indicators (e.g., heart disease). In the present study, we used data from 2 study years (September 1, 2000−August 31, 2001 and September 1, 2001−August 31, 2002), required that persons remain continuously enrolled during each study year, and defined this cohort as the full sample. Subjects were followed each study year from the September 1 start date until their death or August 31, whichever occurred first. Database confounders were captured in the 1-year baseline period before each study year (September 1, 1999−August 31, 2000 and September 1, 2000−August 31, 2001). To make fuller use of available database information in the present study compared with prior studies, we also defined additional database covariates using a broader range of data, including medications, laboratory test results, other health care utilization (e.g., home health services), and disease severity measures, based on methods described previously [11].