ELSEVIER

Journal of Clinical Epidemiology 65 (2012) 384−393

Journal of Clinical Epidemiology

Within-person study designs had lower precision and greater susceptibility to bias because of trends in exposure than cohort and nested case−control designs

Jennifer M. Nicholas*, Andrew P. Grieve, Martin C. Gulliford

Department of Primary Care and Public Health Sciences, King's College London, 7th Floor, Capital House, 42 Weston Street, London SE1 3QD, United Kingdom

Accepted 14 September 2011; Published online 23 December 2011

Abstract

Objective: To compare precision and apparent bias between cohort, nested case−control, self-controlled case series, case−crossover, and case−time−control study designs.

Study Design and Setting: Study designs were implemented to evaluate the association between thiazolidinediones (TZDs) and heart failure, TZDs and fracture, and liver enzyme−inducing anticonvulsants and fracture.

Results: Effect estimates were similar for the cohort and case−control study; for the association between TZDs and fracture in women, the hazard ratio was 1.36 (1.18, 1.56) and odds ratio (OR) was 1.44 (1.21, 1.70). For this clinical example, the self-controlled case series gave upward bias when follow-up was censored at the outcome (incidence rate ratio [IRR], 7.08; 4.96, 10.09) but was otherwise unbiased (IRR, 1.41; 1.14, 1.75). The retrospective case−crossover OR was 3.24 (2.18, 4.80), which was reduced by either bidirectional sampling (OR, 1.20; 0.98, 1.46) or with the case−time−control design (OR, 1.40; 1.09, 1.81). Findings on apparent bias were similar for the other two clinical examples. In each clinical example, within-person designs had considerably lower precision than the cohort or case−control study designs.

Conclusion: When long-term exposures are analyzed, within-person study designs may have lower precision and greater susceptibility to bias. Bias may be reduced by sampling follow-up both before and after the outcome or with the case−time−control study design.

Keywords: Epidemiologic research design; Bone fractures; Heart failure; Thiazolidinediones; Antiepileptics; Drug toxicity

1. Introduction

Epidemiologic research studies, especially those that use secondary data analysis, may test the hypothesis of interest through the selection of one of several potential study design options. Conventional designs, including cohort and nested case−control studies, may yield estimates that are biased by residual confounding because obtaining complete information on all relevant confounders may be difficult [1]. Recently, increasing attention has been paid to within-person study designs, including self-controlled case series and case−crossover designs. Within-person designs are considered to offer improved control over confounding arising from variables that are constant within an individual, such as socioeconomic position, by focusing on comparisons between different periods of time within each individual's follow-up.

Self-controlled case series and case−crossover designs use restricted samples, only including data for participants who experienced the outcome of interest. The self-controlled case series design adopts a cohort perspective through a comparison of the rate of the outcome of interest between exposed and unexposed time periods for each individual [2,3]. Control for risk factors that are constant within an individual is achieved through fitting Poisson regression models, conditional on individual, to estimate the incidence rate ratio (IRR) [2], as in a matched cohort study. The case−crossover design adopts a case−control perspective through a comparison of exposure status between a risk period just before occurrence of the outcome and one or more reference period(s) when the outcome did not occur [4,5]. Control for risk factors that that are constant within an individual is achieved through using analysis methods that condition on matched time periods [4], analogous to a matched case−control study.

* Corresponding author. Tel.: +44-20-7848-6633; fax: +44-20-7848-6620.

E-mail address: jennifer.nicholas@kcl.ac.uk (J.M. Nicholas).

0895-4356/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi: 10.1016/j.jclinepi.2011.09.004

What is new?

- Self-controlled case series, case–crossover, and case–time–control study designs may offer improved control over confounding, but there has been little research into their performance in studies with long-term follow-up.
- This study found that when examining long-term drug exposures, the self-controlled case series and case–crossover study designs may be susceptible to bias from changing exposure probability and may have lower precision than the cohort or nested case–control studies.
- When examining long-term drug exposures, the case–time–control study design may be able to adjust for changes in exposure probability over time but may still have lower precision than the cohort or nested case–control studies.
- Investigators considering the use of within-person designs should carefully evaluate whether conditions that facilitate unbiased estimation using these designs are likely to be met.

The case–crossover and self-controlled case series designs were originally introduced for the study of the acute effects of short-term exposures, such as the occurrence of febrile convulsions after vaccination [3] or triggering of myocardial infarction by recent coffee consumption [4]. However, these designs have been increasingly applied to the study of longer term exposures, such as the use of antipsychotic medications [6]. Application of within-person designs to the analysis of longer term rather than shorter term exposures requires reconsideration of the assumptions underlying these designs.

In both the case–crossover and self-controlled case series designs, adjustment may be required for confounding by variables that vary over time, such as disease severity [2]. In addition, for unbiased estimation, the self-controlled case series design requires that the exposure distribution, and the ability to observe this exposure distribution, is unrelated to event times [7]. Bias may be introduced if the outcome influences the likelihood of future exposures, as when the prescription of a given drug is contraindicated by the outcome event. Bias may also arise if the outcome event leads to censoring, preventing future exposure assessment. Unbiased estimation in the case–crossover design requires that for each participant, the odds of exposure do not vary during follow-up, conditional on variables [8]. The case–time–control design is a modification of the case–crossover design that adjusts for secular trends in exposure by evaluating the change in odds of exposure in a separate sample of control participants who did not experience the outcome [9]. However, the case–time–control study may still be biased if the exposure time trend differs between case and control participants [10,11].

This study aimed to compare the performance of five epidemiologic study designs when used to test the same hypothesis in the same data set. The study designs were cohort, nested case–control, self-controlled case series, case–crossover, and case–time–control designs. Three clinical examples were evaluated using electronic health records from family practices in the United Kingdom. The first example examined the association between thiazolidinediones (TZDs), rosiglitazone, and pioglitazone and heart failure in participants with diabetes. The second example evaluated the association between TZDs and fracture in participants with diabetes. The final example concerned the association of anticonvulsants that induce the cytochrome P450 system [12] with fracture in participants with epilepsy. The liver enzyme–inducing (LEI) anticonvulsants include carbamazepine, oxcarbazepine, phenobarbital, phenobarbital sodium, methylphenobarbital, phenytoin, phenytoin sodium, fosphenytoin sodium, topiramate, and primidone. Effect estimates were compared with reference values derived from published research, and the degree of apparent bias was compared between designs. The precision of effect estimates was also compared between designs.

2. Materials and methods

2.1. Data source and participants

The study designs were implemented using data from the UK General Practice Research Database (GPRD), a large database of anonymized longitudinal electronic medical records from family practices throughout the United Kingdom [13]. In the United Kingdom, 98% of the population is registered with a family practice and a considerable proportion of patient registrations are stable over many years. The GPRD includes information on all medical diagnoses made, and prescriptions issued, by family physicians. Several studies have evaluated the validity of GPRD data with satisfactory results [14,15].

Participants were included in the TZD samples if they were diagnosed with type 2 diabetes mellitus or prescribed oral hypoglycemic drugs or insulin. Participants were excluded if they were diagnosed with type 1 diabetes mellitus, aged younger than 30 years at diabetes onset, or prescribed insulin within 180 days of diabetes onset. Participants with preexisting heart failure were excluded from the TZD and heart failure sample. Follow-up for each participant in the TZD samples started on the later of the date of diabetes onset or July 1, 2000. For the TZD and heart failure example, participant follow-up was censored on diagnosis of heart failure. Participants were included in the LEI anticonvulsant sample if they were diagnosed with epilepsy and prescribed anticonvulsant drugs during follow-up. Follow-up in the LEI anticonvulsant sample started on the later of