

# Performance of prior event rate ratio adjustment method in pharmacoepidemiology: a simulation study<sup>†</sup>

Md Jamal Uddin<sup>1</sup>, Rolf H. H. Groenwold<sup>1,2</sup>, Tjeerd P. van Staa<sup>1,3</sup>, Anthonius de Boer<sup>1</sup>, Svetlana V. Belitser<sup>1</sup>, Arno W. Hoes<sup>2</sup>, Kit C. B. Roes<sup>2</sup> and Olaf H. Klungel<sup>1,2\*</sup>

<sup>1</sup>Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Utrecht, the Netherlands

<sup>2</sup>Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

<sup>3</sup>Health eResearch Centre, Farr Institute for Health Informatics Research, University of Manchester, UK

## ABSTRACT

**Purpose** Prior event rate ratio (PERR) adjustment method has been proposed to control for unmeasured confounding. We aimed to assess the performance of the PERR method in realistic pharmacoepidemiological settings.

**Methods** Simulation studies were performed with varying effects of prior events on the probability of subsequent exposure and post-events, incidence rates, effects of confounders, and rate of mortality/dropout. Exposure effects were estimated using conventional rate ratio (RR) and PERR adjustment method (i.e. ratio of RR post-exposure initiation and RR prior to initiation of exposure).

**Results** In the presence of unmeasured confounding, both conventional and the PERR method may yield biased estimates, but PERR estimates appear generally less biased estimates than the conventional method. However, when prior events strongly influence the probability of subsequent exposure, the exposure effect from the PERR method was more biased than the conventional method. For instance, when the effect of prior events on the exposure was  $RR = 1.60$ , the effect estimate from the PERR method was  $RR = 1.13$  and from the conventional method was  $RR = 2.48$  (true exposure effect,  $RR = 2$ ). In all settings, the variation of the estimates was larger for the PERR method than for the conventional method.

**Conclusion** The PERR adjustment method can be applied to reduce bias as a result of unmeasured confounding. However, only in particular situations, it can completely remove the bias as a result of unmeasured confounding. When applying this method, theoretical justification using available clinical knowledge for assumptions of the PERR method should be provided. Copyright © 2014 John Wiley & Sons, Ltd.

**KEY WORDS**—prior event; post-event; rate ratio; bias; unmeasured confounding; epidemiology; pharmacoepidemiology; simulation

Received 2 February 2014; Revised 25 August 2014; Accepted 22 September 2014

## INTRODUCTION

Unmeasured confounding may impair the validity of observational (pharmaco)epidemiological studies. A recently proposed method, the prior event rate ratio (PERR) adjustment method, may reduce bias as a

result of unmeasured confounding.<sup>1–5</sup> PERR adjustment is a type of self-controlled design in which the exposure effect is estimated by the ratio of two rate ratios (RRs): RR after initiation of exposure ( $RR_{\text{post}}$ ) and the RR prior to initiation of exposure ( $RR_{\text{prior}}$ ).<sup>2,6</sup>

A previous simulation study by Yu *et al.*<sup>1</sup> evaluated and addressed some critical methodological issues of the PERR method. Yu *et al.* mainly focused on different degrees of association between the confounders and the exposure and different effects of the confounders on the outcome in the prior periods and post-periods. They showed that the PERR method can reduce bias as a result of unmeasured confounding when the effect of exposure on the outcome is relatively large in comparison with the interaction effects of exposure and confounders on the outcome or when the time interval effect is rather modest. In addition,

\* Correspondence to: O. H. Klungel, Division Pharmacoepidemiology and Clinical Pharmacology, University of Utrecht, Universiteitsweg 99, PO Box 80082, 3584CG, Utrecht, the Netherlands. Email: O.H.Klungel@uu.nl

<sup>†</sup>The manuscript in this form has not been published and is not being considered for publication elsewhere, in whole or in part, in any language. The abstract of this manuscript has been presented at the Mid-Year Meeting of the International Society for Pharmacoepidemiology, April 5–8, 2014, Rotterdam, the Netherlands, and at the annual epidemiology conference, WEON, June 5–6, 2014, Leiden, the Netherlands. The abstract has also been presented at the 30th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, October 24–27, 2014, Taipei, Taiwan.

this method can reduce bias when unmeasured confounder effects do not vary temporally.<sup>1,2,5</sup>

However, apart from the previously stated situations, the PERR method has not been studied extensively, and the performance of the PERR adjustment method is unclear in several situations, including situations in which there is an influence of prior events on the probability of subsequent exposure and post-events; different incidences of the outcome; different effects of unmeasured confounders, prior events, and exposure on the mortality/dropout in the post-period; and different rates of mortality/dropout. Therefore, further methodological investigations into the PERR adjustment method are required to clarify constraints and understand its proper applicability. Our objective was to assess the performance of the PERR method using simulations under various scenarios and to provide guidelines for application of the method.

## METHODS

We used Monte Carlo simulations to assess the performance of the PERR adjustment method. First, we briefly describe the PERR adjustment method. Second, the simulation set-up, data simulation, and data analysis are described.

### *PERR adjustment method*

The PERR adjustment method can be used in a setting where neither the exposed nor unexposed patients are

treated with the study drugs before the start of the follow-up (Figure 1).<sup>2</sup> In this setting, the RR observed in the prior period, before initiation of the exposure, is because of differences in patient characteristics between the two study groups. The RR in the post-period, after initiation of the exposure, is because of those differences in patient characteristics and exposure. This method requires assumptions about constant temporal effects, that is, confounding effects are constant across prior exposure and post-exposure initiation periods, there is no confounder-by-treatment interaction, and outcomes are non-terminal events.<sup>2,5,7</sup> The exposure effects from the PERR adjustment method is defined as follows:

$$PERR = \frac{\text{Rate ratio during post period}}{\text{Rate ratio during prior period}} = \frac{RR_{post}}{RR_{prior}} \quad (1)$$

Throughout the manuscript, we will use the abbreviation 'PERR' to indicate the exposure effect from the PERR adjustment method.

The prior event rate ratio can be estimated either on incidence RRs or hazard ratios.<sup>3</sup> Here, the PERR was estimated using incidence RRs. Confidence intervals can be obtained by bootstrapping as it is difficult to estimate the covariance between the RRs of the prior periods and post-periods.<sup>1,2,4</sup>

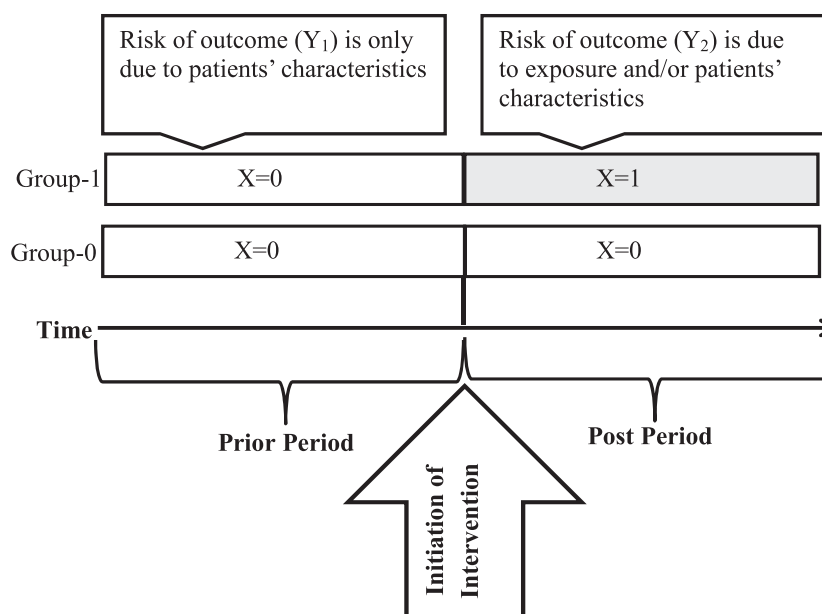


Figure 1. Prior event rate ratio adjustment method.  $X = 1$  and  $X = 0$  represent the exposed and unexposed subjects respectively.  $Y_1$  and  $Y_2$  represent the outcomes in the prior periods and post-periods respectively. The prior period means before intervention, and the post-period means after intervention

### Simulation set-up

The following notation is used:  $X$  denotes the binary exposure/treatment, subscript numbers 1 and 2 indicate prior periods and post-periods, respectively,  $Y_1$  and  $Y_2$  denote the outcomes for prior periods and post-periods,  $C_{11}$  and  $C_{12}$  denote one binary confounder of a subject in prior periods and post-periods respectively (for example,  $C_{11}$  and  $C_{12}$  are a measurement of blood pressure of a patient in the prior periods and post-periods respectively), and  $M_2$  denotes the mortality/dropout in the post-period. Although we considered binary confounders, the performance of PERR is expected to be similar in the case of continuous confounders.<sup>1</sup>

We considered several scenarios for the simulation study, which are graphically depicted in causal diagrams. The following scenarios can be identified. Scenario 1: There is confounding in both periods and confounders ( $C_{11}$  and  $C_{12}$ ) are mutually associated as they are from the same patient in the prior periods and post-periods. We assessed the PERR when the confounder effects are constant across both periods. We also evaluated robustness of the method to violation of the assumption of constant effects of confounders in both periods. In that case, the confounder effects are varied between prior periods and post-periods, for example, no differences (effects of  $C_{11}$  and  $C_{12}$  are constant), small-to-moderate differences (e.g. the effect of  $C_{11}$  on  $X$  and  $Y_1$  is RR 1.5, and the effect of  $C_{12}$  on  $X$  and  $Y_2$  is RR 2.0), and profound differences (e.g. the effect of  $C_{11}$  on  $X$  and  $Y_1$  is RR 2.0, and the effect of  $C_{12}$  on  $X$

and  $Y_2$  is RR 3.5). In this setting, the prior events do not influence the exposure and post-events (Figure 2a). In all other scenarios (i.e. scenarios 2–4), the confounders are associated and effects of confounders are constant across the periods. Scenario 2: Prior events directly influence the post-events but do not influence the exposure (Figure 3a). Scenario 3: Prior events influence the probability of subsequent exposure in the absence and presence of confounders (Figures 4a and 4c respectively). Furthermore, the prior events do not directly influence the post-events. Scenario 4: There is mortality/dropout in the post-period that is influenced by the confounder ( $C_{12}$ ), exposure ( $X$ ), and the prior event ( $Y_1$ ). In addition, the prior events do not influence the exposure and the post-events (Figure 5a).

We also examined the impact of different incidences of the outcomes on the PERR and interaction effects of exposure and confounders ( $C_{12}$ ) on the outcomes.

In each simulation, the sample size was 100 000 and each scenario was replicated for 10 000 times. The incidence rate of the outcome was varied between 1 and 10% in the prior period and 1 and 20% in the post-period. The prevalence of exposure was varied between 20 and 25%, and the mortality/dropout rate was varied between 1 and 20%.

### Data simulation

Data were simulated as follows. The prior period confounder ( $C_{11}$ ) follows a binomial distribution with mean probability between 0.20 and 0.30. As the PERR

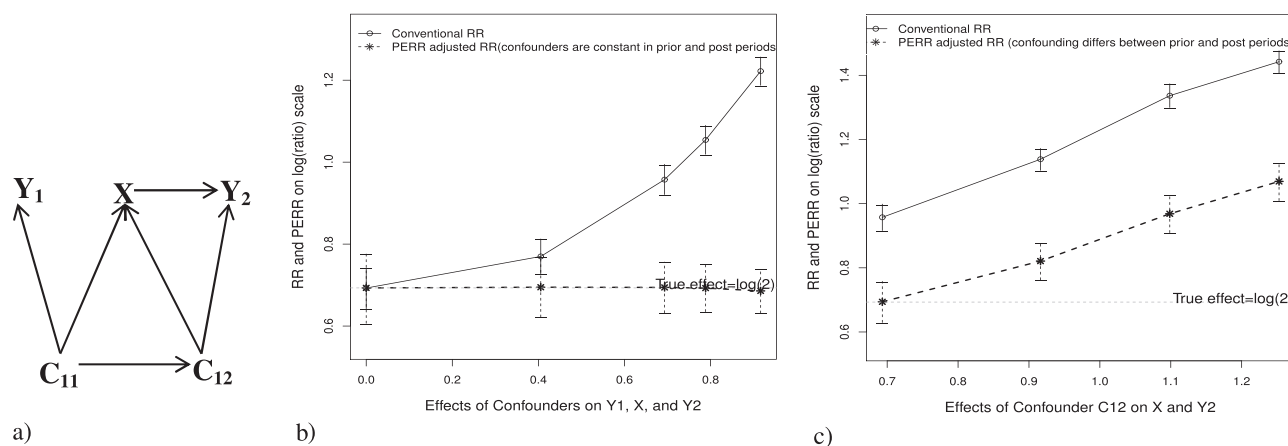


Figure 2. Impact of confounders on estimate adjusted by prior event rate ratio and conventional estimate. (a) Directed acyclic graph that illustrates that there are confounders in the both periods (either constant or different) and confounders are associated. Nevertheless, the prior events ( $Y_1$ ) do not influence the probability of subsequent exposure ( $X$ ) and the post-events ( $Y_2$ ).  $X$  represents the exposure,  $Y_1$  and  $Y_2$  represent the events in the prior periods and post-periods, respectively, and  $C_{11}$  and  $C_{12}$  represent a confounder in the prior periods and post-periods, respectively. (b) Effects of confounders on PERR and conventional RR when effects of confounders are constant between prior periods and post-periods. (c) Effects of confounders on PERR and conventional RR when the effect of confounders is different between prior periods and post-periods. RR—rate ratio and PERR—prior event rate ratio. The dotted horizontal straight line (b–c) represents the true exposure effect (RR = 2.00,  $\log(2) = 0.6931$ ). Vertical bars indicate the range for 95% of the estimates and are based on the 2.5 and 97.5 percentiles of the distribution of the estimates. Results are based on simulations with sample size of 100 000, and each scenario was simulated for 10 000 times

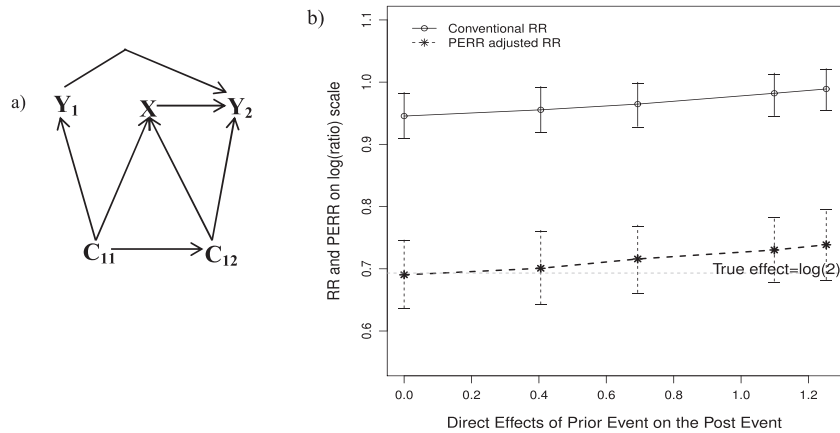


Figure 3. Estimate adjusted by prior event rate ratio when prior events ( $Y_1$ ) directly influence the post-events ( $Y_2$ ). (a) Directed acyclic graph that illustrates that there are confounders in both periods and confounders are associated and prior events directly influence the post-events. However, prior event did not influence the exposure.  $X$  represents the exposure,  $Y_1$  and  $Y_2$  represent the events in the prior periods and post-periods respectively;  $C_{11}$  and  $C_{12}$  represent a confounder in the prior periods and post-periods respectively. (b) Effects of prior events ( $Y_1$ ) on the post-events ( $Y_2$ ). RR—rate ratio and PERR—prior event rate ratio. The dotted horizontal straight line represents the true exposure effect (RR = 2.00,  $\log(2) = 0.6931$ ). Vertical bars indicate the range for 95% of the estimates and are based on the 2.5 and 97.5 percentiles of the distribution of the estimates. Results are based on simulations with sample size of 100 000, and each scenario was simulated for 10 000 times

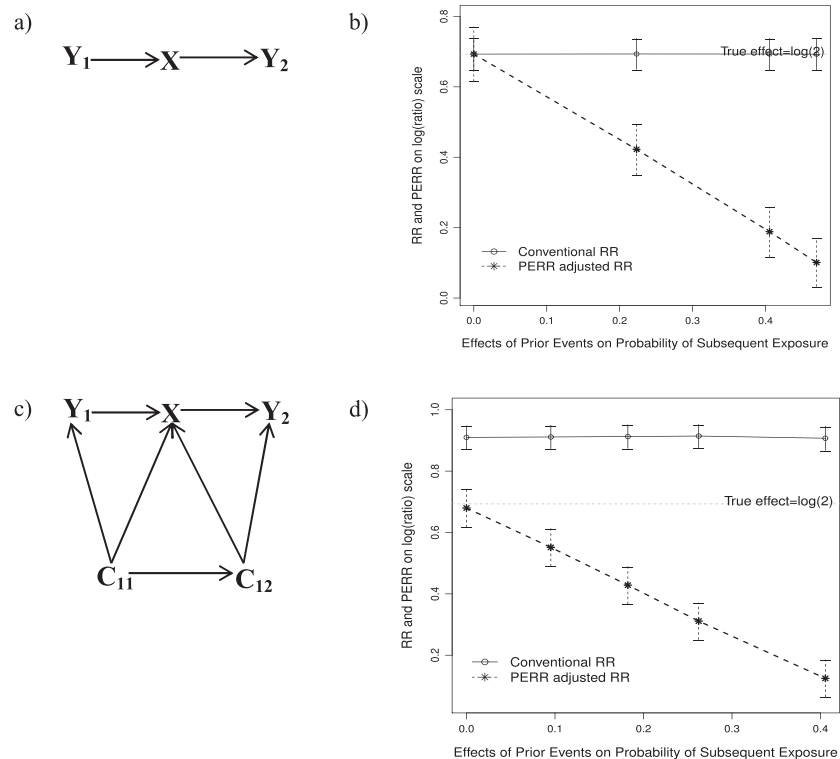


Figure 4. Estimate adjusted by prior event rate ratio when prior events ( $Y_1$ ) influence the probability of subsequent exposure ( $X$ ). (a) Directed acyclic graph (DAG) that illustrates that there is no confounding variable in both periods, but prior events ( $Y_1$ ) influence the probability of subsequent exposure ( $X$ ). (b) Effects of 'prior' events on the probability of subsequent exposure in the absence of confounders. (c) DAG that illustrates that there are confounders in both periods, confounders are associated, and the 'prior' events influence the probability of subsequent exposure. (d) Effects of 'prior' events on the probability of subsequent exposure in the presence of confounders.  $X$  represents the exposure,  $Y_1$  and  $Y_2$  represent the events in the prior periods and post-periods respectively, and  $C_{11}$  and  $C_{12}$  represent a confounder in the prior periods and post-periods respectively. RR—rate ratio and PERR—prior event rate ratio. The dotted horizontal straight line (b and d) represents the true exposure effect (RR = 2.00,  $\log(2) = 0.6931$ ). Vertical bars indicate the range for 95% of the estimates and are based on the 2.5 and 97.5 percentiles of the distribution of the estimates. Results are based on simulations with sample size of 100 000, and each scenario was simulated for 10 000 times

# PERFORMANCE OF PERR ADJUSTMENT METHOD

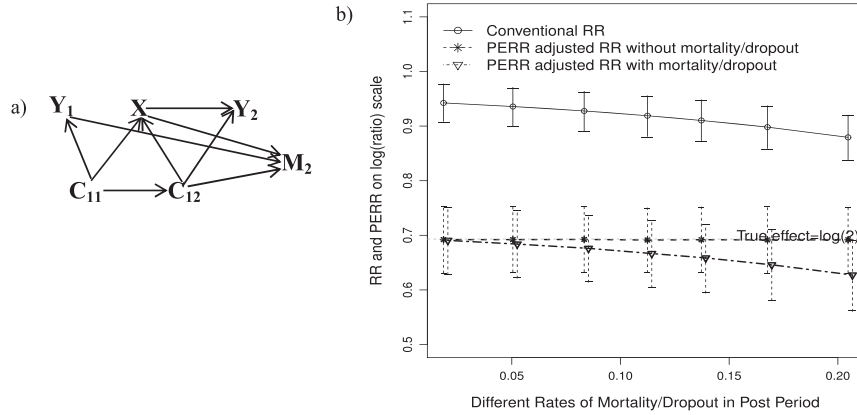


Figure 5. Estimate adjusted by prior event rate ratio and conventional estimate when there is mortality or dropout. (a) Directed acyclic graph that illustrates that there is mortality or dropout ( $M_2$ ) that is influenced by the confounders ( $C_{12}$ ), prior events ( $Y_1$ ), and exposure ( $X$ ); confounders in both periods are associated and constant; the prior event does not influence the probability of subsequent exposure.  $X$  represents the exposure,  $Y_1$  and  $Y_2$  represent the outcome in the prior periods and post-periods respectively, and  $C_{11}$  and  $C_{12}$  represent a confounder in the prior period and post-period respectively;  $M_2$  represents the mortality or dropout in the post-period. (b) Impact of mortality or dropout on the PERR and conventional RR. RR—rate ratio and PERR—prior event rate ratio. The dotted horizontal straight line represents the true exposure effect (RR = 2.00, log(2) = 0.6931). Vertical bars indicate the range for 95% of the estimates and are based on the 2.5 and 97.5 percentiles of the distribution of the estimates. Results are based on simulations with sample size of 100 000, and each scenario was simulated for 10 000 times. The mortality/dropout rate is about 5–20%

method estimates the ‘marginal’ effect of exposure on the outcome, we used log-linear models in order to generate our data, to prevent issues of non-collapsibility, because the log-linear model is collapsible.<sup>8,9</sup> In addition, we assumed that the time period is fixed and incidence rates are constant across periods, and hence, the estimated RR is equivalent to the risk ratio, which is also collapsible.<sup>10,11</sup>

The log-linear models were used for generating prior events (Equation 2), confounder ( $C_{12}$ ) in the post-period (Equation 3), exposure (Equation 4), post-events (Equation 5), and mortality or dropout (Equation 6). The models are described in the succeeding text:

$$P_{y_1} = \exp(\alpha_{y_1} + \beta_{c_{11}.y_1} C_{11}) \text{ and } Y_1 \sim \text{Bernoulli}(P_{y_1}) \quad (2)$$

$$P_{c_{12}} = \exp(\alpha_{c_{12}} + \beta_{c_{11}.c_{12}} C_{11} + \beta_{y_1.c_{12}} Y_1) \text{ and } C_{12} \sim \text{Bernoulli}(P_{c_{12}}) \quad (3)$$

$$P_x = \exp(\alpha_x + \beta_{c_{11}.x} C_{11} + \beta_{c_{12}.x} C_{12} + \beta_{y_1.x} Y_1) \text{ and } X \sim \text{Bernoulli}(P_x) \quad (4)$$

$$P_{y_2} = \exp(\alpha_{y_2} + \beta_{c_{12}.y_2} C_{12} + \beta_{x.y_2} X + \beta_{y_1.y_2} Y_1) \text{ and } Y_2 \sim \text{Bernoulli}(P_{y_2}) \quad (5)$$

$$P_{m_2} = \exp(\alpha_{m_2} + \beta_{c_{12}.m_2} C_{12} + \beta_{x.m_2} X + \beta_{y_1.m_2} Y_1) \text{ and } M_2 \sim \text{Bernoulli}(P_{m_2}) \quad (6)$$

where  $P_{y_1}$  and  $P_{y_2}$  are the probability of the outcome in the prior periods and post-periods respectively.  $P_x$ ,  $P_{m_2}$ , and  $P_{c_{12}}$  are the probability of the exposure, mortality, and confounders ( $C_{12}$ ) respectively.  $\beta_{c_{11}.x}$ ,  $\beta_{c_{11}.y_1}$  and  $\beta_{c_{11}.c_{12}}$  are the effects of prior period confounders ( $C_{11}$ ) on the exposure, prior events, and the post-period confounders respectively, and values of these parameters were varied between log(1) and log(10).  $\beta_{y_1.c_{12}}$ ,  $\beta_{y_1.x}$ ,  $\beta_{y_1.y_2}$ , and  $\beta_{y_1.m_2}$  are the effects of prior events on the post-period confounders ( $C_{12}$ ), exposure, post-events, and mortality/dropout respectively, and values of these parameters were varied between log(1) and log(5).  $\beta_{c_{12}.x}$ ,  $\beta_{c_{12}.y_2}$ , and  $\beta_{c_{12}.m_2}$  are the effects of post-period confounders ( $C_{12}$ ) on the exposure, post-events, and mortality/dropout respectively, and values of these parameters were varied between log(1) and log(5).  $\beta_{x.y_2}$  and  $\beta_{x.m_2}$  are the effects of exposure on the post-events and mortality/dropout respectively, and the values of these parameters were varied between log(1) and log(2).  $\alpha_{y_1}$  and  $\alpha_{y_2}$  denote the intercepts of the prior and post-outcome models respectively, which indicate baseline incidence rates.  $\alpha_x$  denotes the probability (20%) of the exposure among those without prior events and confounder values of zero.  $\alpha_{c_{12}}$  and  $\alpha_{m_2}$  are the intercepts for the  $P_{c_{12}}$  and  $P_{m_2}$ . To ensure that the probabilities of  $P_{c_{12}}$  and  $P_{m_2}$  were between 0 and 1, the values of the parameters  $\alpha_{c_{12}}$  and  $\alpha_{m_2}$  were varied between  $-5$  and  $0$ .

## Data analysis

The PERR adjustment method aims to control for unmeasured confounding. Therefore, in all analyses, the



confounders ( $C_{11}$  and  $C_{12}$ ) were considered unmeasured. The PERR adjustment exposure effect was estimated using Equation 7a, in the case of mortality/dropout, and Equation 7b, without mortality/dropout below. The conventional RR was estimated using the data from the post-period only, again omitting the (unmeasured) confounder  $C_{12}$  from the model. We also analysed data using the log-linear model.

The estimation formula of the PERR is as follows:

$$\begin{aligned} PERR &= \frac{\text{Rate ratio during post period}}{\text{Rate ratio during prior period}} \\ &= \frac{E(Y_2|X=1 \& M_2=0)/E(Y_2|X=0 \& M_2=0)}{E(Y_1|X=1)/E(Y_1|X=0)} \end{aligned} \quad (7a)$$

where  $M_2=0$  denotes whether a subject remained alive in the post-period. When there is no mortality/dropout, Equation 7a can be written as

$$PERR = \frac{E(Y_2|X=1)/E(Y_2|X=0)}{E(Y_1|X=1)/E(Y_1|X=0)} \quad (7b)$$

For the scenario in which the prior events influence the probability of subsequent exposure (i.e. scenario 3), we also estimated the PERR using an adjustment of the observed prior events by inverse probability weighting using the propensity score.

The exposure effects from the PERR and the conventional method were estimated in each simulated dataset, log-transformed, and averaged across datasets. The variation of the estimates was estimated by the standard deviation of the 10 000 estimates. We also estimated 95% confidence intervals in a non-parametric way using the 2.5 and 97.5 percentiles of the 10 000 estimates. We used the statistical software package R (Windows, version 2.15.1) to simulate and analyse the data.<sup>12</sup>

## RESULTS

Results based on Scenario 1 (in which confounders in both periods were associated, either constant or different across prior periods and post-periods, and the prior events did not influence the probability of subsequent exposure and the post-events) showed that the conventional RR was biased, RR 2.16 to 4.23 (true RR=2.00), when the association between confounders and exposure and confounders and outcome was RR 1.50 to 3.5 respectively. In this setting, the association between the confounders in the prior periods and post-periods was RR=10. When the effects of confounders

are constant across prior periods and post-periods, the PERR without mortality/dropout was unbiased (Figure 2b). However, when the effects of confounders are not constant, the PERR was biased (RR 2.27 to 2.91), just as the conventional analysis, but the bias was smaller than that for conventional analysis (Figure 2c). The bias for conventional RR and PERR was similar when the confounders in both periods were independent (data not shown).

Figure 3b shows the results for Scenario 2, in which prior events directly influence the post-events but not the subsequent exposure. The PERR was biased, and the bias increased with increasing the relation between prior events and post-events. For example, the PERR was 2.04 to 2.10 (true RR=2.00) when the effect of prior events on post-events was RR 1.50 to 3.50 respectively. Moreover, the conventional RR was biased, but the magnitude of bias was much larger than that of the PERR method (Figure 3b).

Figure 4 shows the results for Scenario 3, in which prior events influence the probability of subsequent exposure in the absence and presence of confounders in both periods respectively. In the first case, the PERR was highly biased, and the bias increased with increasing the effects of prior events on the probability of exposure (Figure 4b). For example, the PERR was 1.53 to 1.11 (true RR=2.00) when the effect of prior events on exposure was RR 1.25 to 1.60 respectively. However, the conventional RR was unbiased. Similarly, in the second case, the PERR was also highly biased, and the bias increased as the effects of prior events increased (Figure 4d). In this case, the conventional RR was also biased, RR=2.49 (true RR=2.00). When the prior events strongly influence the probability of the exposure (e.g. RR=1.60), the bias was more pronounced for PERR (PERR=1.13) than for the conventional RR (RR=2.48; Figure 4d). Unlike the bias for PERR in Figure 2c where bias was because of unmeasured confounding and the direction of bias was positive (overestimate the effect), in this case, the direction of bias for PERR was negative (underestimate the effect). In Scenario 3 with a null exposure effect (RR=1.00), the PERR shifted away from the null. In a separate analysis of Scenario 3, the PERR was estimated using an adjustment of the observed prior events by inverse probability weighting. This analysis showed that the estimates of the PERR method were similar to that of the conventional method (Table 1).

Figure 5 shows the results of Scenario 4, in which we assessed the impact of an effect of confounders, exposure, and prior events on mortality/dropout. The PERR with taking mortality/dropout into account was

# PERFORMANCE OF PERR ADJUSTMENT METHOD

Table 1. Comparison between conventional rate ratio (RR) and both propensity-score-adjusted PERR (using an adjustment of observed prior event by inverse probability weighting) and unadjusted PERR

Effects of prior event on subsequent exposure in the absence of confounding effect and prior event does not influence the post-event (i.e. Scenario 3, Figure 4a)	PERR (unadjusted; in log-ratio scale)	PERR (propensity score adjusted; in log-ratio scale)	Conventional RR (in log-ratio scale)
log(1.00)	0.6970	0.6909	0.6902
log(1.25)	0.4387	0.6935	0.6940
log(1.50)	0.2046	0.6913	0.6905
log(1.60)	0.1367	0.6920	0.6920
Effects of prior event on subsequent exposure in the presence of confounding effects, but prior event does not influence the post-event (i.e. Scenario 3, Figure 4c)	PERR (unadjusted; in log-ratio scale)	PERR (propensity score adjusted; in log-ratio scale)	Conventional RR (in log-ratio scale)
log(1.00)	0.6831	0.9121	0.9125
log(1.10)	0.5512	0.9094	0.9098
log(1.20)	0.4308	0.9128	0.9145
log(1.50)	0.2249	0.8830	0.8871
log(1.60)	0.1399	0.8801	0.8838

Results are based on simulations with sample size of 100 000, and each scenario was simulated for 10 000 times. The true exposure effect is  $RR = 2$  ( $\log(RR) = 0.6931$ ). PERR, prior event rate ratio.

more biased (PERR 1.90, rate of mortality = 17% and true  $RR = 2.00$ ) than the PERR without taking mortality/dropout into account (PERR = 2.00). The bias increased with increasing rates of mortality/dropout.

In all scenarios, the variation of the estimates was more pronounced for PERR than for the conventional RR (Table 2). For example, when the effect of prior events on the exposure was  $RR = 1.50$ , the standard deviations of the conventional RR and PERR were 0.020 and 0.031 respectively. The variation of the estimates decreased with increasing incidence rates. For instance, when incidence rates were 1 and 20% in the post-period, the standard deviations of the PERR were 0.080 and 0.023 respectively.

We also simulated and analysed data in a situation where interaction effects of exposure and unmeasured confounders were present; we observed that the interaction effects induced bias in the PERR. When prior events did not influence the probability of exposure and post-events, and the impact of unmeasured confounding was constant between prior periods and post-periods, and there was no interaction between

exposure and unmeasured confounder, PERR showed unbiased results.

## DISCUSSION

Our simulation study shows that in the presence of unmeasured confounding, the PERR adjustment method and conventional regression analysis both yield biased estimates of the treatment effect. However, the PERR method results in less biased estimates than the conventional method when confounding differs considerably between prior periods and post-periods or when prior events directly influence the post-events or when there is mortality or dropout. When prior events influence the probability of the subsequent exposure (e.g.  $RR \geq 1.20$ ), the bias is more pronounced for the PERR method than for the conventional method. Moreover, the bias increased with increasing strength of the relation between prior events and exposure.

In our simulations, the magnitude of bias was smaller for the PERR method than for the conventional method when the effects of the unmeasured confounder differed

Table 2. Comparison between standard deviation of conventional rate ratio (RR) and the PERR-adjusted RR

Different effects of prior events on the exposure in the presence of confounders	Conventional RR		PERR-adjusted RR	
	Estimate (in log-ratio scale)	Standard deviation *	Estimate (in log-ratio scale)	Standard deviation *
log( $RR = 1.00$ )	0.9100	0.0195	0.6798	0.0319
log( $RR = 1.10$ )	0.9113	0.0195	0.5515	0.0310
log( $RR = 1.20$ )	0.9127	0.0194	0.4284	0.0309
log( $RR = 1.30$ )	0.9143	0.0194	0.3112	0.0303
log( $RR = 1.50$ )	0.9072	0.0199	0.1246	0.0308

Results are based on simulations with sample size of 100 000, and each scenario was simulated for 10 000 times.

\*Standard deviation of the 10 000 estimates. The true exposure effect is  $RR = 2$  ( $\log(2) = 0.6931$ ).

between the prior periods and post-periods (i.e. violation of the assumption of constant confounding effect in both periods), but again, both methods were biased. This is an important finding because the PERR method has been proposed as a method to deal with unmeasured confounding.<sup>1-5</sup> If the assumption of constant confounding effects is violated, the PERR method may thus yield biased estimates.

The PERR that is estimated by considering mortality/dropout was more biased than those of PERR without taking mortality/dropout into account. The magnitude of bias increases as the rate of mortality/dropout increases. We found that the interaction effects of exposure and unmeasured confounder induced bias in the PERR, which is in line with the study of Yu *et al.*<sup>1</sup> Yu *et al.* evaluated both PERR and an alternative formulation of the PERR (PERR-ALT), but we only focused on the PERR. The PERR-ALT is a ratio of two RRs: the RR between the post-periods and prior periods for the exposed group and the RR between the post-periods and prior periods for the unexposed group.<sup>1</sup> Because we ignored a set-up of clustering in our data and covariates in the analysis, both PERR and PERR-ALT methods provide similar results.

In all settings, the variation of the estimates was larger for PERR than for the conventional RR. This may be because of the fact that PERR is the ratio of two RRs between the prior periods and the post-periods. The variation decreases as incidence rates increase in both periods. Tannen *et al.*<sup>2</sup> stated that the variation in the PERR can be larger with a smaller number of prior events. We argue that if researchers believe that the conventional point estimates are substantially biased because of unmeasured time-fixed confounding, the PERR adjustment method can be a viable alternative if assumptions are met. Furthermore, because the PERR method yields large standard errors compared with conventional methods, the method can be considered more conservative than conventional methods.

We observed that both the magnitude of bias and its direction vary with the effects of prior events on the probability of exposure as well as the different effects of confounders between prior periods and post-periods. For a given effect of prior event and confounders, the bias may be either negative or positive, which is related to the ratio of the two RRs. More clearly, if the RR before initiation of the exposure is larger than the RR after initiation of the exposure, the denominator of the PERR is larger than the numerator; the resulting exposure effect is moving to the null effect and bias becomes negative. In that case, when

prior event strongly influences the probability of the subsequent exposure, the exposure effect is very close to the null effect. A similar explanation applies to positive bias when bias is because of the unmeasured confounders.

The major limitation of the PERR method is the requirement that the event rate prior to study start can be estimated. Thus, this method cannot be applied when patients have no records prior to the study start or for patients who die prior to exposure initiation.<sup>1,4</sup> Apart from these limitations, the findings of our simulation study suggest that the PERR method is not an adequate method to remove bias as a result of unmeasured confounding in situations when prior events influence the probability of subsequent exposure even when there is no exposure effect on the outcome. An example of likely bias with PERR would be the study of statin use and risk of myocardial infarction (MI) or stroke. When a patient first experienced an MI or stroke, she or he has to start treatment (e.g. statin) in order to avoid a secondary MI or stroke, and consequently, the health characteristics of this patient (e.g. cholesterol levels) are influenced by such type of therapy. Hence, the probability of the exposure (start statin therapy) is definitely influenced by the previous outcome as well as the confounders.

The PERR method may be a valid method to remove bias as a result of unmeasured confounding in a situation where the probability of exposure and the probability of post-events are independent of the prior events. An example may be a study of the effect of proton-pump inhibitors (PPIs) on the risk of pneumonia. Patients may experience pneumonia for multiple times in their life; thus, it can occur before and/or after initiation of the PPIs; hence, one of the prerequisites to apply the PERR method (outcome before intervention) is met. Moreover, if a patient has suffered pneumonia before initiation of the PPIs, this prior outcome may not influence the likelihood of being prescribed PPIs because PPIs are generally prescribed to relieve symptoms of gastroesophageal reflux disease and not of pneumonia. Hence, another important condition (independence between prior event and subsequent exposure) of the PERR adjustment method is satisfied here. However, in this setting, the PERR method may provide biased results when the unmeasured confounder effects vary temporally in prior periods and post-periods, and prior events directly influence the post-events. Another example where the PERR method could also be valid is a study of statin use and risk of hip fracture or lower urinary tract symptoms. Like in the PPI example, this example may also satisfy the critical assumptions underlying



the PERR method. However, if the rate of mortality or dropout is high, or prior events influence the post-events, or the unmeasured confounder effects vary temporally, the PERR adjustment method may not provide unbiased results.

We note that the PERR adjustment method has some similarities with the case-time-control (CTC) design and case-crossover design (CCD).<sup>13–15</sup> For example, all three methods deal with intermittent exposures with transient effects and overcome confounding by constant characteristics.<sup>1,13,15,16</sup> However, there are some differences between these methods. Differences between the CCD and the PERR methods are that in the PERR method, both cases and non-cases are used, whereas CCD only uses information on cases. Furthermore, in the CCD, exposure misclassification is more likely, and bias may arise in the selection of the control time window(s)<sup>15</sup> but not in the PERR method. In addition, the PERR method generally deals with time-to-event outcome, but CCD deals with binary outcomes and is analysed using conditional logistic regression to account for the matched nature of the data.<sup>17</sup> A possible limitation of the CCD is that it does not account for general time trend in drug use,<sup>13</sup> while this limitation does not appear in the PERR. Differences between the CTC and the PERR methods include that CTC also deals with binary outcome, but the PERR generally deals with time-to-event outcome.<sup>1,13</sup> The CTC assumes common period effect for cases and controls, whereas the PERR does not make this assumption.<sup>1,13</sup>

For the simplest situation in Figure 2, the prior event  $Y_1$  acts just like a negative control (i.e. an outcome such that the set of unmeasured confounders  $C_{12}$  of exposure  $X$  and post-event  $Y_2$  should be as identical as possible to the set of unmeasured confounders  $C_{11}$  of exposure  $X$  and prior event  $Y_1$ )<sup>18</sup> for detecting unmeasured confounding. If the negative control is empirically associated with the exposure after adjustment for possible measured confounders, unmeasured confounding may be present in the data.<sup>18–21</sup> The prior event also acts as a ‘perturbation variable’ (a variable that is associated with the exposure and post-event only through unmeasured confounders) to detect and correct (via adjustment of the perturbation variable) the bias of unmeasured confounding.<sup>18–20</sup> Detailed explanations of negative controls and perturbation variables can be found in Lipsitch *et al.*,<sup>18</sup> Flanders *et al.*,<sup>20</sup> Lee,<sup>19</sup> and Tchetgen Tchetgen.<sup>21</sup>

For a complex scenario, such as when prior events influence the probability of subsequent exposure (i.e. Scenario 3, Figure 4), one can adjust for the prior events (e.g. via inverse probability weighting using propensity score). Once this adjustment is performed,

the PERR method yields the same result as the conventional method. Although it is easy to assess such complex scenarios in simulations, this may be very hard in empirical studies. Therefore, before going to apply the PERR adjustment method in empirical data, theoretical motivations on the appropriateness of applying the method should be given based on clinical knowledge.

Although Yu *et al.*<sup>1</sup> indeed pointed out some crucial situations for the PERR adjustment method, our simulations go beyond those and showed that in particular cases (e.g. the prior event rate strongly influences the probability of exposure), the PERR method provides more biased estimates than the conventional method. Moreover, we show that the PERR method is sensitive to small violation of one of the key assumptions (i.e. constant effects of confounders in both prior periods and post-periods). In addition, we simulated and analysed data in the case of dropout and mortality, which is important in a follow-up study.

Strengths of our simulation studies are that we generated data in several scenarios that are commonly seen in pharmacoepidemiological research; the conventional RR and PERR are estimated by a simple ratio technique, which is easy to understand and free from different assumptions of statistical modelling techniques, and the sample size and number of replications are large enough to achieve enough power. Limitations of our study are that the time effects and incidence rates are constant across prior periods and post-periods. Future research could explore settings of different time interval effects and different incidence rates in both periods. Although we did not estimate coverage probabilities of the PERR, we expect that the coverage probabilities are far away from the nominal level (especially in the reported scenarios) because of a large amount of bias for the PERR method.

In conclusion, the PERR adjustment method can be applied to reduce bias as a result of unmeasured confounding. However, in particular situations, for example, when prior events strongly influence the probability of subsequent exposure, this method can be more biased than conventional methods. Hence, when applying this method, we suggest providing theoretical justification using available clinical knowledge for underlying assumptions of the PERR method.

## CONFLICT OF INTEREST

Olaf Klungel had received unrestricted funding for pharmacoepidemiological research from the Dutch private–public funded Top Institute Pharma (TI Pharma Grant T6.101 Mondriaan).

## KEY POINTS

- Extensive simulation studies were performed to evaluate the performance of the prior event rate ratio (PERR) adjustment method in several real-istic (pharmaco)epidemiological settings.
- In the presence of unmeasured confounding, both conventional analysis and the PERR method may yield biased estimates.
- In particular situations, (when unmeasured con-founding differs considerably between prior and post-periods or when prior events directly influ-ence the post-events or when there is mortality or dropout), the PERR method yields less biased estimates than the conventional method. How-ever, when prior events strongly influence the probability of subsequent exposure, PERR can be more biased than the conventional method.
- Variation of the estimates is larger for the PERR method than for the conventional methods.

## ACKNOWLEDGEMENTS

The research leading to these results was conducted as part of the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT consortium), which is a public–private partnership coordi-nated by the European Medicines Agency.

The PROTECT project is supported by Innovative Medicine Initiative (IMI) Joint Undertaking ([www.imi.europa.eu](http://www.imi.europa.eu)) under Grant Agreement no 115004, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and European Federation of Pharma-ceutical Industries and Associations companies' in-kind contribution. In the context of the IMI Joint Undertaking, the Department of Pharmacoepidemiology, Utrecht Uni-versity, also received a direct financial contribution from Pfizer. The views expressed are those of the authors only and not of their respective institution or company.

## REFERENCES

1. Yu M, Xie D, Wang X, Weiner MG, Tannen RL. Prior event rate ratio adjustment: numerical studies of a statistical method to address unrecognized confounding in observational studies. *Pharmacoepidemiol Drug Saf* 2012; **21**: 60–68.
2. Tannen RL, Weiner MG, Xie D. Use of primary care electronic medical record database in drug efficacy research on cardiovascular outcomes: comparison of database and randomised controlled trial findings. *BMJ (Online)* 2009; **338**: 395–399.
3. Tannen R, Xie D, Wang X, Yu M, Weiner MG. A new "Comparative Effective-ness" assessment strategy using the THIN database: comparison of the cardiac complications of pioglitazone and rosiglitazone. *Pharmacoepidemiol Drug Saf* 2013; **22**: 86–97.
4. Tannen RL, Weiner MG, Xie D. Replicated studies of two randomized trials of angiotensin-converting enzyme inhibitors: further empiric validation of the 'prior event rate ratio' to adjust for unmeasured confounding by indication. *Pharmacoepidemiol Drug Saf* 2008; **17**: 671–685.
5. Weiner MG, Xie D, Tannen RL. Replication of the Scandinavian Simvastatin Survival Study using a primary care medical record database prompted explora-tion of a new method to address unmeasured confounding. *Pharmacoepidemiol Drug Saf* 2008; **17**: 661–670.
6. Schneeweiss S, Seeger JD, Smith SR. Methods for developing and analyzing clinically rich data for patient-centered outcomes research: an overview. *Pharmacoepidemiol Drug Saf* 2012; **21**: 1–5.
7. Alemayehu D, Alvir JM, Jones B, Willke RJ. Statistical issues with the analysis of nonrandomized studies in comparative effectiveness research. *J Manage Care Pharm* 2011; **17**: S22–S26.
8. Agresti A. Categorical data analysis. John Wiley and Sons: Hoboken, New Jersey, 2002.
9. Ali MS, Groenwold RHH, Pestman WR, *et al.*. Propensity score balance measures in pharmacoepidemiology: a simulation study. *Pharmacoepidemiol Drug Saf* 2014; **23**: 802–811.
10. Rothman KJ. Epidemiology: an introduction. Oxford University Press: New York, USA, 2012.
11. Greenland S. Confounding of incidence density ratio in case-control studies. *Epidemiology* 2013; **24**: 624–625.
12. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing Vienna Austria 2010.
13. Suissa S. The case-time-control design. *Epidemiology* 1995; **6**: 248–253.
14. Suissa S, Dell'Aniello S, Martinez C. The multitime case-control design for time-varying exposures. *Epidemiology* 2010; **21**: 876–883.
15. Schneeweiss S, Stürmer T, Maclure M. Case-crossover and case-time-control de-signs as alternatives in pharmacoepidemiologic research. *Pharmacoepidemiol Drug Saf* 1997; **6**: S51–S59.
16. Maclure M, Mittleman AM. Should we use a case-crossover design? *Annu Rev Public Health* 2000; **21**: 193–221.
17. Delaney JA, Suissa S. The case-crossover study design in pharmacoepidemiology. *Stat Methods Med Res* 2009; **18**: 53–65.
18. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* 2010; **21**: 383–388.
19. Lee W. Detecting and correcting the bias of unmeasured factors using perturbation analysis: a data-mining approach. *BMC Med Res Methodol* 2014; **14**.
20. Flanders WD, Klein M, Darrow LA, *et al.*. A method for detection of residual confounding in time-series and other observational studies. *Epidemiology* 2011; **22**: 59–67.
21. Tchetgen Tchetgen E. The control outcome calibration approach for causal infer-ence with unobserved confounding. *Am J Epidemiol* 2014; **179**: 633–640.