PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2016; **25**: 287–296
Published online 16 December 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.3924

ORIGINAL REPORT

Conditions for confounding of interactions†

Aihua Liu[1,2,3], Michal Abrahamowicz[1,2]* and Jack Siemiatycki[3,4]

[1]*Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada*
[2]*Division of Clinical Epidemiology, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada*
[3]*Division of Population Health, CRCHUM Research Center, Montreal, Quebec, Canada*
[4]*Department of Social and Preventive Medicine, University of Montreal, Montreal, Quebec, Canada*

ABSTRACT

Purpose Pharmaco-epidemiology increasingly investigates drug–drug or drug–covariate interactions. Yet, conditions for confounding of interactions have not been elucidated. We explored the conditions under which the estimates of interactions in logistic regression are affected by confounding bias.

Methods We rely on analytical derivations to investigate the conditions and then use simulations to confirm our analytical results and to quantify the impact of selected parameters on the bias of the interaction estimates.

Results Failure to adjust for a risk factor U results in a biased estimate of the interaction between exposures $E1$ and $E2$ on a binary outcome Y if the association between U and $E1$ varies depending on the value of $E2$. The resulting confounding bias increases with increase in the following: (i) prevalence of confounder U; (ii) strength of U–Y association; and (iii) heterogeneity in the association of $E1$ with U across the strata of $E2$. A variable that is not a confounder for the main effects of $E1$ and $E2$ may still act as an important confounder for their interaction.

Conclusions Studies of interactions should attempt to identify—as potential confounders—those risk factors whose associations with one of the exposures in the interaction term may be modified by the other exposure. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS—interaction; effect modification; confounding bias; logistic regression; simulations; pharmacoepidemiology

Received 24 February 2015; Revised 16 October 2015; Accepted 28 October 2015

Pharmaco-epidemiology increasingly investigates drug–drug interactions[1–4] or interactions between drug exposure and some users' characteristics.[5,6] Yet, interaction assessment is methodologically challenging.[7–13] Knol *et al*. reported considerable between-study variation regarding the methods to assess interaction and present the results; although they discussed several methodological issues, they did not discuss the issue of potential confounding of interaction estimates.[13] Similarly, other methodological papers on interaction and subgroup analyses did not discuss the conditions for unbiased estimation[8–11] or even mention the issue.[14] This is surprising, given the paramount importance of confounding of main effects in epidemiological research.[15]

We screened 30 articles that investigated interactions, published in leading epidemiological and clinical journals: only six mentioned potential confounding bias, and all of these relied on concepts borrowed from assessment of confounding of main effects.[16–22] However, the conditions for confounding of main effects may not necessarily apply to confounding of interaction between *two* exposures.

This paper relies on both algebraic derivations and simulations to explore conditions under which a failure to adjust for a variable induces confounding bias in the estimated effect of an interaction between two other variables in logistic regression. In this article, we adapt a pragmatic definition of confounding as a systematic difference between the estimate of the relationship of interest and the corresponding true parameter, which is due to a failure to account for other variable(s) but cannot be explained merely by the noncollapsibility of the odds ratios (ORs). In other words, while noncollapsibility implies that the logistic regression-based ORs are biased toward the null

*Correspondence to: M. Abrahamowicz, Division of Clinical Epidemiology, McGill University Health Centre, Royal Victoria Hospital, 687 Pine Avenue West, V-Pavilion, V2.20A, Montreal, QC, H3A 1A1, Canada. E-mail: michal.abrahamowicz@mcgill.ca

†Prior posting and presentations: Part of the results was presented in an oral presentation of the 3rd North American Congress of Epidemiology, Montreal, 21–24 June 2011.

whenever the model fails to adjust for an important risk factor,[16,23–25] the presence, direction, and strength of confounding depend also on the relationships between this risk factor and the variables whose effects are being estimated.[26,27]

CONDITIONS FOR CONFOUNDING OF INTERACTION

Basic definitions and concepts

We consider the inter-relations of four variables: $E1$ and $E2$ are two binary exposures, whose potential interaction for a binary outcome Y is investigated, and U is a binary covariate, a potential confounder of the interaction. Our analytical and simulation results, presented in this manuscript, apply to any exposure or treatment, and any covariates that satisfy the earlier assumptions.

The main goal of interaction analyses may be to assess either the joint effect of two exposures or whether the effect of a single exposure is modified by another variable. Joint effects are of interest if, for example, $E1$ and $E2$ represent use of two specific drugs, and the $E1 * E2$ interaction is tested primarily to assess if their effects are independent of each other (no interaction) or synergistic. Effect modification analyses may investigate, for example, whether, and how, the effect of exposure to the drug represented by $E1$ varies depending on the value of some binary characteristic such as sex, represented by E2. While there are conceptual differences between these two paradigms, in practice, both rely on the valid estimation of, and statistical inference about, the parameter for the $E1 * E2$ interaction. Our analytical and simulation results apply to both types of interaction analyses.

We assume that U is not on the causal pathway between either $E1$ or $E2$ and Y. In our notation: OR_{E1-Y} refers to the OR for the marginal association between $E1$ and Y, which ignores a possible interaction with $E2$; $\mathrm{OR}_{E1-Y|E2=x}$ refers to OR_{E1-Y} conditional on $E2 = x$, and $\mathrm{ROR}_{(\mathrm{E1*E2})-Y}$ refers to the effect of the $E1 * E2$ interaction for Y, which can be quantified as the ratio of two ORs:[28,29]

$$\mathrm{ROR}_{(E1*E2)-Y} = \frac{\mathrm{OR}_{E1-Y|E2=1}}{\mathrm{OR}_{E1-Y|E2=0}} \qquad (1)$$

$\mathrm{ROR}_{(E1*E2)-\mathrm{Y}}$ equals $\exp(\beta)$, where β is the regression coefficient for the interaction $E1 * E2$ in the logistic regression model, which also includes regression coefficients for $E1$ and $E2$.

If U is a confounder for $\mathrm{ROR}_{(E1*E2)-Y}$, then the estimate *not* adjusted for U will differ systematically from its true value, implying the following inequality:

$$R_{\mathrm{Inter}} = \frac{\text{the estimated } \mathrm{ROR}_{(E1*E2)-Y} \text{ while not adjusting for } U}{\text{the true } \mathrm{ROR}_{(E1*E2)-Y}} \neq 1 \qquad (2)$$

In the following, we explore the conditions under which $R_{\mathrm{Inter}} \neq 1$, that is, the conditions necessary for U to act as a confounder for $\mathrm{ROR}_{(E1*E2)-Y}$.

Necessary and sufficient conditions for U to be a confounder of $ROR_{(E1*E2)-Y}$

In Appendices A–C, we demonstrate that each of the following conditions is necessary for $R_{\mathrm{Inter}} \neq 1$ in Equation 2:

(A): U is associated with Y: $\mathrm{OR}_{U-Y} \neq 1$;

(B): U is associated with each of $E1$ and $E2$, in at least one stratum of the other exposure ($E2$ or $E1$): $[\mathrm{OR}_{E1-\mathrm{U}|E2=0} \neq 1$ or $\mathrm{OR}_{E1-U|E2=1} \neq 1]$ and $[\mathrm{OR}_{E2-\mathrm{U}|E1=0} \neq 1$ or $\mathrm{OR}_{E2-\mathrm{U}|E1=1} \neq 1]$;

(C): The association between $E1$ and U varies across the strata of $E2$: $\mathrm{OR}_{E1-U|E2=0} \neq \mathrm{OR}_{E1-U|E2=1}$.

While proving the preceding conditions, we refer to Table 1, which presents frequencies (number of subjects in a hypothetical population) for each of 16 subgroups, corresponding to different combinations of Y, $E1$, $E2$, and U. For Y, 0 and 1 correspond to non-diseased and diseased, respectively; for other variables, 0 and 1 correspond to unexposed and exposed, respectively. When U is *not* adjusted for, the 16 subgroups are collapsed into eight subgroups, by merging the corresponding subgroups with $U = 0$ and $U = 1$ (last column of Table 1).

OR_{U-Y} is assumed constant across combinations of $E1$ and $E2$, implying no interactions of U with either $E1$ or $E2$ for Y. As shown in Appendix A, R_{Inter} in Equation 2 can be written in terms of the frequencies in Table 1:

$$R_{\mathrm{Inter}} = \frac{1+\frac{b1}{a1}}{1+\frac{b1}{a1\mathrm{OR}_{U-Y}}} * \frac{1+\frac{d1}{c1\mathrm{OR}_{U-Y}}}{1+\frac{d1}{c1}} * \frac{1+\frac{f1}{e1\mathrm{OR}_{U-Y}}}{1+\frac{f1}{e1}} * \frac{1+\frac{h1}{g1}}{1+\frac{h1}{g1\mathrm{OR}_{U-Y}}} \qquad (3)$$

Clearly, whenever $\mathrm{OR}_{U-Y} = 1$, $R_{\mathrm{Inter}} = 1$. Thus, condition A ($\mathrm{OR}_{U-Y} \neq 1$) is a necessary condition for $R_{\mathrm{Inter}} \neq 1$.

In Appendix B, we demonstrate that conditions B and C are also necessary conditions for $R_{\mathrm{Inter}} \neq 1$.

Given that condition C logically implies B, the conjunction of conditions A and C constitutes a sufficient condition for $R_{\mathrm{Inter}} \neq 1$ in Equation 2.

DOI: 10.1002/pds

Table 1. Frequencies of the 16 hypothetical subgroups

				Frequencies	
Y*	$E1$†	$E2$†	U‡	When U is adjusted for§	When U is not adjusted for¶
0	0	0	0	$a0$	$a0 + b0$
0	0	0	1	$b0$	
1	0	0	0	$a1$	$a1 + b1$
1	0	0	1	$b1$	
0	1	0	0	$c0$	$c0 + d0$
0	1	0	1	$d0$	
1	1	0	0	$c1$	$c1 + d1$
1	1	0	1	$d1$	
0	0	1	0	$e0$	$e0 + f0$
0	0	1	1	$f0$	
1	0	1	0	$e1$	$e1 + f1$
1	0	1	1	$f1$	
0	1	1	0	$g0$	$g0 + h0$
0	1	1	1	$h0$	
1	1	1	0	$g1$	$g1 + h1$
1	1	1	1	h1	

*Y is the binary outcome.
†$E1$ and $E2$ are two binary exposure variables.
‡U is a potential binary "confounder."
§Frequencies in this column correspond to the expected frequency of subjects with specific values of Y, $E1$, $E2$, and U, which would be observed if the data were analyzed with the logistic model: $\text{logit}(\text{Prob}[Y=1]) = \beta_0 + \beta_{E1}E1 + \beta_{E2}E2 + \beta_U U + \beta_{\text{Int}}(E1 * E2)$.
¶Frequencies in this column correspond to the expected frequency of subjects with specific values of Y, $E1$, and $E2$, which would be observed if the data were analyzed with the logistic model that excludes U: $\text{logit}(\text{Prob}[Y=1]) = \beta_0 + \beta_{E1}E1 + \beta_{E2}E2 + \beta_{\text{Int}}(E1 * E2)$.

Appendix C shows that a failure to account for U will also result in a biased estimate of $\text{ROR}_{(E1*E2)-Y}$, in the following conditions: (i) if U is associated with $E1$ and (ii) if U modifies the effect of $E2$ on Y (implying interaction between U and $E2$ for Y).

Some distortions of ORs may be due to their noncollapsibility.[30] To demonstrate that the conditions we outline pertain to the concept of confounding rather than noncollapsibility, we have carried out analogous derivations while focusing on relative risk (RR) rather than OR. Appendix D shows that the right hand of Equation 3 represents the ratio of the interaction estimate not adjusted for U to the corresponding true parameter for the $E1 * E2$ for Y interaction, regardless of whether RRs or ORs are used as the effect measure. Thus, conditions A and C apply also to RRs and identify the situations where failure to adjust for U will result in estimated RR, for $E1 * E2$ interaction, differing from its true value. Because RR is not affected by noncollapsiblility,[23,31,32] such difference will mostly reflect the impact of model mis-specification due to omission of a confounder.

EMPIRICAL EXAMPLE

To illustrate the practical usefulness of our condition C, we used data from a case–control study of lung cancer conducted in 1996–2002 in Montreal.[33,34] A total of 1211 newly diagnosed primary lung cancer cases (743 men and 468 women) were frequency matched on sex and age with 1539 population controls (923 men and 616 women). Information on socioeconomic status, ethnic group, cigarette smoking history, and other risk factors for lung cancer was collected. Details of study design and data collection were reported elsewhere.[33]

We focus on the differences in lung cancer risk between men and women and explore two examples of potential effect measure modifications that may be of substantive interest in etiology of lung cancer. Specifically, we investigate, in separate analyses, if the following modify the association between sex and lung cancer risk: (i) ethnicity and (ii) income. These analyses imply testing interactions between sex ($E1$: exposure of primary interest) and either ethnicity or income, representing a potential effect modifier ($E2$). Ethnicity was dichotomized as French Canadian (two-thirds of participants) versus all others, and income as above versus below the median income of all subjects.

The current analyses were limited to 730 women and 1426 men who were past or current smokers and had complete data. We considered a binary indicator of "heavy smoking" (above the median of the lifetime cigarette pack-years among all subjects) as the potential unmeasured confounder (U). Indeed, while smoking is the main risk factor for lung cancer, it is not available in large administrative databases. Each of the two interaction effects was analyzed separately and was

DOI: 10.1002/pds

estimated using two different unconditional multivariable logistic regression models, which, respectively, did or did not adjust for smoking (*U*). We then assessed if failure to adjust for smoking affected the estimated ROR for each interaction.

Table 2 compares the RORs, for each of the two interactions, estimated with and without adjustment for heavy smoking. Failure to adjust for smoking did the following: (1) changed substantially the estimated ROR for the sex * ethnicity interaction but (2) had almost no impact on the ROR for the sex * income interaction (Table 2). To explain this contrast, the rightmost column of Table 2 shows that the association between heavy smoking and sex (1) varies significantly by ethnicity ($ROR_{(Sex*Ethnicity)\text{-}Smoking} = 1.90$, 95%CI: 1.20, 3.00) but (2) does *not* depend on income ($ROR_{(Sex*Income)\text{-}Smoking} = 0.98$, 95%CI: 0.66, 1.43). Thus, smoking does meet our condition C in the analyses focusing on the sex * ethnicity interaction for lung cancer but does not meet this condition for the sex * income interaction. Given that condition C is necessary for the confounding of the interaction effect estimate, the preceding contrast explains why failure to adjust for smoking (which clearly meets our condition A) affects only the sex * ethnicity interaction. This contrast also provides further support for our conclusion that the observed change for sex * ethnicity interaction is due to unmeasured confounding rather than noncollapsibility. Indeed, if the impact of not adjusting for smoking reflected noncollapsibility, then it should do the following: (i) affect both interactions to a similar extent and (ii) lead to the estimated interaction being weaker (with coefficient closer to 0) than the estimate obtained while conditioning on smoking. Yet, results in Table 2 show that neither of these consequences occurred.

SIMULATION STUDIES

Objectives

First, simulations were designed to confirm the analytical results and to assess how the bias of the point estimate of $ROR_{(E1*E2)-Y}$ depended on the following: (i) the joint distribution of *U*, *E*1, and *E*2; (ii) the effect of *U* on *Y*; and (c) the prevalence of *U*. Secondly, simulations allowed us to account for sampling error and to assess the impact of confounding on the statistical inference about the estimated interaction, as well as its impact on the estimated effects of *E*1 in specific strata of *E*2, and vice versa.

Simulation design and data generation methods

We simulated a hypothetical study, with $N = 2000$ subjects, of the interaction between two binary exposures *E*1 and *E*2, for binary outcome (*Y*), with a continuous covariate *C*, correlated with *E*1. A binary risk factor (*U*), associated with *Y*, was not adjusted for in the analyses. Different scenarios assumed different patterns of the relationships of *U* with *E*1 and *E*2 and with *Y*. Appendix E describes distributions of all variables and their relationships, assumed in different simulated scenarios.

Twenty-two scenarios corresponding to different combinations of selected parameters were simulated (see left parts of Tables 3 and 4). Table 3 summarizes scenarios designed to corroborate the results of analytical derivations. Table 4 presents scenarios that assessed the impact of selected parameters.

The binary outcome *Y* was generated from the multivariable logistic regression model:

Table 2. Empirical example of the impact of un-adjusted for risk factor on the estimated RORs for interactions between other variables (data derived from the Montreal case–control study of lung cancer, 1996–2002, $N = 2156$)

		$ROR_{(E1*E2)\text{-Lung cancer}}$*†					
		Adjusting for smoking history‡		Not adjusting for smoking history§		$ROR_{(E1*E2)\text{-smoking}}$¶	
*E*1	*E*2	Point estimate	95%CI	Point estimate	95%CI	Point estimate	95%CI
Sex	Ethnicity**	1.47	0.93, 2.33	1.73	1.12, 2.67	1.90	1.20, 3.00
Sex	Income†††	1.28	0.87, 1.88	1.22	0.85, 1.76	0.98	0.66, 1.43

Abbreviations: CI = confidence interval; ROR = exponentiated coefficient for the interaction.
*Multivariable logistic regression model with binary outcome defined as lung cancer cases versus controls, including the main effects of sex, ethnicity, age, and income.
†See Table I in Appendix I for the estimates of the regression coefficients of primary interest.
‡The model included the main effects of sex, age, ethnicity/income, and the binary variable for heavy smoking versus non-heavy smoking.
§The model included the main effects of sex, age, and ethnicity or income.
¶Binary outcome was defined as heavy smoking (pack-years ≥ median value of pack-years for all subjects) versus non-heavy smoking.
**Ethnicity was coded as French Canadian versus all other ethnicity groups.
††Income was coded as high income (income ≥ median value of income for all subjects) versus non-high income.

DOI: 10.1002/pds

Table 3. Simulation scenarios for assessing necessary and sufficient conditions for U to be a confounder of the $E1 * E2$ interaction for Y

	Scenario assumptions*						Results from the models that do not adjust for U				
	Probability of $U=1$ (%) for subgroup with										
Scenario no.	$E1=1$, $E2=1$	$E1=1$, $E2=0$	$E1=0$, $E2=1$	$E1=0$, $E2=0$	β_U†	β_{Int}‡	$\overline{\beta}_{\text{Int}}$§	R_{Inter}¶	Relative bias (%) **	Coverage rate of 95%CI (%)	% of $p<0.05$††
1	40.0	10.0	30.0	12.5	0.0	0.0	0.02	1.04		96.0	4.0
2	40.0	10.0	30.0	12.5	0.0	−0.4	−0.39	1.03	−2.5	96.4	53.4
3	40.0	40.0	20.0	20.0	0.7	0.0	0.02	1.03		96.0	4.0
4	40.0	40.0	20.0	20.0	0.7	−0.4	−0.38	1.03	−5.0	95.8	52.3
5	40.0	20.0	30.0	13.8	0.7	0.0	0.05	1.06		95.7	4.4
6	40.0	20.0	30.0	13.8	0.7	−0.4	−0.36	1.06	−10.0	95.8	45.0
7	40.0	10.0	30.0	12.5	0.7	0.0	0.11	1.13		92.6	7.4‡‡
8	40.0	10.0	30.0	12.5	0.7	−0.4	−0.30	1.13	−25.0	93.3	32.0

Abbreviations: CI = confidence interval; OR = odds ratio.
*In all the scenarios listed in Table 3, the following values of $\beta = \ln(\text{OR})$ are assumed for binary exposures $E1$ and $E2$ and the adjusted for continuous covariate C: $\beta_{E1}=0.4$, $\beta_{E2}=0.9$, and $\beta_C=0.02$.
†Value of $\ln(\text{OR}_{U-Y})$ for the effect of U on Y, assumed for the corresponding scenario.
‡True value of $\ln[\text{ROR}_{(E1*E2)-Y}]$ for the interaction between $E1$ and $E2$ for Y, assumed for the corresponding scenario.
§Mean of 2000 estimated $\ln[\text{ROR}_{(E1*E2)-Y}]$ for the interaction between $E1$ and $E2$ for Y.
¶The mean value of R_{Inter}, defined in Equation 2, calculated as $\exp(\overline{\beta}_{\text{Int}})/\exp(\beta_{\text{Int}})$.
**Calculated as $(\overline{\beta}_{\text{Int}}-\beta_{\text{Int}})/(\beta_{\text{Int}})$ (%).
††The percentages in the last column correspond to either (i) type I error rate, in scenarios where $\beta_{\text{Int}}=0$ or (ii) empirical power of the interaction test of $\beta_{\text{Int}}\neq 0$.
‡‡Inflated type I error rate; that is, the 95%CI for type I error exceeds 0.05.

$$\begin{aligned}\text{logit}(\text{Prob}[Y=1]) &= \beta_0+\beta_{E1}E1+\beta_{E2}E2\\ &\quad+\beta_C C+\beta_U U\\ &\quad+\beta_{\text{Int}}(E1{*}E2)\end{aligned} \tag{4}$$

The intercept (β_0) was selected to obtain approximately 50% "disease" ($Y=1$) prevalence. In all scenarios, we assumed $\beta_{E1}=0.4$, $\beta_{E2}=0.9$, and $\beta_C=0.02$, while varying β_U and β_{Int} (Tables 3 and 4).

For each scenario, we generated 2000 independent random samples.[35]

Analysis of simulated data

Simulated data were analyzed with multivariable logistic models. All the models included $E1$, $E2$, and C, as well as the $E1 * E2$ interaction (a spurious effect in scenarios with true $\text{ROR}_{(E1*E2)-Y}=1$). Regression coefficient estimates, standard errors, and p-values for two-tailed Wald tests were computed for each sample and summarized across the 2000 simulated samples.

We focused on the estimation of and inference about $\beta_{\text{Int}}=\ln[\text{ROR}_{(E1*E2)-Y}]$, for the $E1 * E2$ interaction. We assessed bias, coverage rate of 95%CI, and type I error or power of the interaction test (see Appendix E for details).

Simulation Results

*Necessary and sufficient conditions for U to confound the E1 * E2 interaction for Y.* Whenever U had no effect on Y, failure to adjust for U did *not* induce bias of the interaction estimates (scenarios 1 and 2 in Table 3), confirming that condition A ($\text{OR}_{U-Y}\neq 1$) is necessary for U to be a confounder of $\text{ROR}_{(E1*E2)-Y}$. Thus, in all other simulated scenarios, we assumed $\text{OR}_{U-Y}\neq 1$.

In scenarios 3 and 4, where U was associated with $E1$ but was not related with $E2$ in either stratum of E1, failure to adjust for U did not bias the $E1 * E2$ interaction estimates (Table 3). Furthermore, Table H in Appendix H shows that the results for scenarios 3 and 4 remain unchanged if the U–Y association becomes much stronger, with OR increasing from 2.0 to 5.0. These results confirm that condition B is also necessary for U to be a confounder of $\text{ROR}_{(E1*E2)-Y}$.

In scenarios 5–8, U had independent associations with both $E1$ and $E2$. However, scenarios 7 and 8 assumed the $E1 * E2$ interaction for U ($\text{ROR}_{(E1*E2)-U}=2$), which—in contrast—was absent in scenarios 5 and 6 ($\text{ROR}_{(E1*E2)-U}=1$). Interestingly, the models that did not adjust for U yielded reasonably accurate results for the $E1 * E2$ interaction for Y in scenarios 5 and 6. In contrast, the corresponding estimates had a relative bias of up to −25.0% in scenarios 7 and 8, where the association $E1$–Y varied across strata of E2, with $\text{ROR}_{(E1*E2)-U}=2$ (Table 3). This contrast confirms that condition C ($\text{OR}_{E1-U|E2=0}\neq\text{OR}_{E1-U|E2=1}$) is a necessary condition for U to be a confounder of the $E1 * E2$ interaction for Y. Furthermore, scenario 7, which assumed no true $E1 * E2$ interaction, shows an inflated type I error

DOI: 10.1002/pds

Table 4. Simulation scenarios for assessing how selected parameters affect the presence/magnitude of confounding bias

	Scenario assumptions*						Results from the models that do not adjust for U				
Scenario no.	β_U†	Prevalence of $U=1$ (%)‡	$\text{ROR}_{(E1*E2)-U}$§	OR_{E1-U}¶	OR_{E2-U}**	β_{Int}††	$\overline{\beta}_{\text{Int}}$‡‡	R_{Inter}§§	Relative bias (%)¶¶	Coverage rate of 95%CI (%)	% of $p<0.05$***
7	0.70	24.6	2.00	1.61	4.62	0.0	0.11	1.13		92.6	7.4†††
7.1	0.70	26.0	4.00	3.16	6.19	0.0	0.23	1.28		76.4	20.7†††
7.2	0.70	28.3	0.50	1.61	3.63	0.0	−0.04	0.98		96.7	3.4
7.3	0.70	22.4	0.25	0.54	2.34	0.0	−0.18	0.85		86.4	13.7†††
8	0.70	24.6	2.00	1.61	4.62	−0.4	−0.30	1.13	−25.0	93.3	32.0
8.1	0.70	26.0	4.00	3.16	6.19	−0.4	−0.18	1.27	−55.0	81.2	13.0
8.2	0.70	28.3	0.50	1.61	3.63	−0.4	−0.44	0.98	10.0	95.8	64.4
8.3	0.70	22.4	0.25	0.54	2.34	−0.4	−0.58	0.85	45.0	86.8	88.4
7	0.70	24.6	2.00	1.61	4.62	0.0	0.11	1.13		92.6	7.4†††
7.4	0.18	24.6	2.00	1.61	4.62	0.0	0.05	1.07		95.4	4.6
8	0.70	24.6	2.00	1.61	4.62	−0.4	−0.30	1.13	−25.0	93.3	32.0
8.4	0.18	24.6	2.00	1.61	4.62	−0.4	−0.37	1.05	−7.5	96.1	49.1
7	0.70	24.6	2.00	1.61	4.62	0.0	0.11	1.13		92.6	7.4†††
7.5	0.70	12.3	2.00	1.61	4.62	0.0	0.07	1.09		95.0	5.0
8	0.70	24.6	2.00	1.61	4.62	−0.4	−0.30	1.13	−25.0	93.3	32.0
8.5	0.70	12.3	2.00	1.61	4.62	−0.4	−0.34	1.08	−15.0	95.3	42.4
7	0.70	24.6	2.00	1.61	4.62	0.0	0.11	1.13		92.6	7.4†††
7.6	0.70	27.2	2.00	1.00	1.07	0.0	0.12	1.15		91.5	8.6†††
7.7	0.70	25.2	4.00	1.00	1.14	0.0	0.19	1.23		84.8	15.2†††
8	0.70	24.6	2.00	1.61	4.62	−0.4	−0.30	1.13	−25.0	93.3	32.0
8.6	0.70	27.2	2.00	1.00	1.07	−0.4	−0.28	1.15	−30.0	91.8	31.3
8.7	0.70	25.2	4.00	1.00	1.14	−0.4	−0.22	1.22	−45.0	86.1	18.4

Abbreviations: CI = confidence interval; OR = odds ratio.

*In all the scenarios listed in Table 4, the following values of $\beta = \ln(\text{OR})$ are assumed for binary exposures $E1$ and $E2$ and the adjusted for continuous covariate C: $\beta_{E1} = 0.4$, $\beta_{E2} = 0.9$, and $\beta_C = 0.02$.

†Value of $\ln(\text{OR}_{U-Y})$ for the effect of U on Y, assumed for the corresponding scenario.

‡Overall prevalence of $U = 1$ across all combinations of $E1$, $E2$, and Y.

§OR of the interaction between $E1$ and $E2$ for U, assumed for the corresponding scenario.

¶OR for the marginal association between $E1$ and U across the two strata of $E2$, assumed for the corresponding scenario.

**OR for the marginal association between $E2$ and U across the two strata of E1, assumed for the corresponding scenario.

††True value of $\ln[\text{ROR}_{(E1*E2)-Y}]$ for the interaction between $E1$ and $E2$ for Y, assumed for the corresponding scenario.

‡‡Mean of 2000 estimated $\ln[\text{ROR}_{(E1*E2)-Y}]$ for the interaction between $E1$ and $E2$ for Y.

§§The mean value of R_{Inter}, defined in Equation 2, calculated as $\exp(\overline{\beta}_{\text{Int}})/\exp(\beta_{\text{Int}})$.

¶¶Calculated as $(\overline{\beta}_{\text{Int}} - \beta_{\text{Int}})/(\beta_{\text{Int}})$ (%).

***The percentages in the last column correspond to either (i) type I error rate, in scenarios where $\beta_{\text{Int}} = 0$, or (ii) empirical power of the interaction test of $\beta_{\text{Int}} \neq 0$.

†††Inflated type I error rate; that is, the 95%CI for type I error exceeds 0.05.

rate (Table 3). This indicates that omitting a variable that meets our conditions for a confounder of the interaction of interest will result in not only biased interaction estimates but also substantial increase in the risk of detecting a spurious interaction when it is truly absent.

Appendix F presents similar results and the same conclusions, based on additional simulations where (i) $E1$ and $E2$ represented continuous, rather than binary, exposures and (ii) U was continuous. Finally, Appendix G shows that the results do not depend on whether $E1$ and $E2$ are correlated or not.

Assessing the impact of selected parameters. Scenarios 7.1–7.7 and 8.1–8.7 differ from, respectively, scenario 7 or 8 on selected parameter(s) (Table 4).

Scenarios 7.1–7.3 and 8.1–8.3 assessed the impact of varying the degree to which the association $E1$–U varied across strata of $E2$ ($\text{ROR}_{(E1*E2)-U}$). The bias of $\beta_{\text{Int}} = \ln[\text{ROR}_{(E1*E2)-Y}]$ estimate due to failing to adjust for U increased as the $E1*E2$ interaction for U became stronger: for example, from a relative bias of −25% with $\text{ROR}_{(E1*E2)-U} = 2$ (scenario 8) to −55% with $\text{ROR}_{(E1*E2)-U} = 4$ (scenario 8.1). Furthermore, the direction of the $E1*E2$ interaction for U determined the direction of the bias in $\text{ROR}_{(E1*E2)-Y}$. Because $\beta_U > 0$ (Table 4), when $\text{ROR}_{(E1*E2)-U} > 1$, the mean estimate for the coefficient for the $E1*E2$ interaction for Y is positively biased, that is, higher than the true value, whereas when $\text{ROR}_{(E1*E2)-U} < 1$, the bias is negative (Table 4).

In scenarios 7.4, 7.5, 8.4, and 8.5, both a weaker effect of U on Y (scenarios 7.4 and 8.4) and a lower

DOI: 10.1002/pds

prevalence of U (scenarios 7.5 and 8.5) resulted in less biased ln[$ROR_{(E1*E2)-Y}$] estimates (Table 4).

Table 5 illustrates the impact of the confounding bias in the estimated coefficient for the $E1 * E2$ interaction on the estimated effects of each of the two exposures, across the strata of the other exposure, for selected scenarios of Table 4. Specifically, the mean estimate of the effect of $E1$ in a given stratum of $E2$ was calculated as the mean value, across 2000 simulations, of $\beta_{E1} + \beta_{Int}E2$, where β_{E1} and β_{Int} are estimates from the logistic model.[9] The resulting mean estimates (shown in the right part of Table 5) were then compared with the corresponding true effects (left part of Table 5). Regardless of whether the interaction is truly absent (upper part of Table 5) or present (lower part), the stratified effects of both exposures differ from their corresponding true effects, with different effects either underestimated or overestimated. For example, when the true effect of $E1$ does not depend on $E2$ (true $\beta_{Int} = 0$), unmeasured confounding due to failure to adjust for U induces spurious differences between the estimated effects of $E1$ in the two strata of $E2$, and *vice versa* (scenarios 7–7.3).

Conditions for confounding differ for interaction effects versus main effects. Scenarios 7.6 and 8.6 assume a combination of the following: (i) no marginal associations of U with either $E1$ or $E2$ (across the two strata of, respectively, E2 and E1) and (ii) a moderately strong $E1 * E2$ interaction for U with $ROR_{(E1*E2)-U} = 2$. Notice that (i) implies that U is *not* a confounder for the main effects of either E1 or E2. The resulting estimates of $ROR_{(E1*E2)-Y}$ were biased and very similar to, respectively, scenarios 7 and 8, where U had strong marginal associations with $E1$ and $E2$ (Table 4). Moreover, in scenarios 7.7 and 8.7, where $ROR_{(E1*E2)-U}$ was increased to 4, the bias became much stronger (Table 4). These results emphasize the importance of our condition C: whenever a risk factor U meets the condition $ROR_{(E1*E2)-U} \neq 1$, it will act as a confounder of $ROR_{(E1*E2)-Y}$ even if, in bivariate analyses, U has no marginal correlations with either E1 or E2. In conclusion, a variable that does *not* meet classic conditions for a confounder of the main effect of either exposure may still act as an important confounder of their interaction.

To further emphasize the contrast between the conditions for confounding of (i) main effects versus (ii) interaction effects, we reanalyzed data from scenarios 7.6 and 7.7, focusing on the estimates of the main effects of $E1$ and $E2$ on Y. Specifically, we estimated the reduced multivariable logistic model that included only the (mutually adjusted) main effects of $E1$ and $E2$ on Y and compared the estimates with the corresponding true effects. Notice that because scenarios 7.6 and 7.7 assume no interaction between $E1$ and $E2$ for Y (true $\beta_{Int} = 0$ in Table 4), the "reduced" main effects model is here very similar to the true data-generating model, except that it does not include U. In all these simulations, the main effects of both exposures were estimated without any systematic bias (relative biases not exceeding 2.5%), even if in the same scenarios the interaction effects showed substantial bias (scenarios 7.6 and 7.7 in Table 4). These results further demonstrate that even if U is *not* a confounder for the main effects of $E1$ and $E2$ on Y, it may be still an important confounder for the $E1 * E2$ interaction for Y, as long as it meets our condition C: $ROR_{(E1*E2)-U} \neq 1$.

DISCUSSION

Any observational (pharmaco-)epidemiologic study is vulnerable to unmeasured confounding bias. Yet, the conditions for confounding of interactions have hardly been discussed in the literature. We have derived the

Table 5. Impact of confounding of the estimated $E1 * E2$ interaction on the bias of the estimated effects (log ORs) of $E1$ in different strata of $E2$, and *vice versa* (in selected simulated scenarios)

Scenario no.*	Scenario assumptions†					Results from the models that do not adjust for U‡				
	$\beta_{E1\|E2=0}$	$\beta_{E1\|E2=1}$	$\beta_{E2\|E1=0}$	$\beta_{E2\|E1=1}$	β_{Int}	$\overline{\beta}_{Int}$	$\beta_{E1\|E2=0}$	$\beta_{E1\|E2=1}$	$\beta_{E2\|E1=0}$	$\beta_{E2\|E1=1}$
7	0.40	0.40	0.90	0.90	0.00	0.11	0.37	0.48	0.99	1.10
7.1	0.40	0.40	0.90	0.90	0.00	0.23	0.39	0.62	0.94	1.17
7.3	0.40	0.40	0.90	0.90	0.00	−0.18	0.38	0.20	1.10	0.92
8	0.40	0.00	0.90	0.50	−0.40	−0.30	0.38	0.08	0.99	0.69
8.1	0.40	0.00	0.90	0.50	−0.40	−0.18	0.39	0.21	0.94	0.76
8.3	0.40	0.00	0.90	0.50	−0.40	−0.58	0.39	−0.19	1.11	0.53

*Details of each scenario are shown in Table 4.
†True value of ln[$ROR_{(E1*E2)-Y}$] for the interaction between $E1$ and $E2$ for Y as well as true values of ln(OR) for the stratified effects of $E1$ and $E2$, assumed for the corresponding scenario.
‡Mean of 2000 estimated ln[$ROR_{(E1*E2)-Y}$] for the interaction between $E1$ and $E2$ for Y and estimated ln(OR) for the stratified effects of $E1$ and $E2$.

DOI: 10.1002/pds

necessary and sufficient conditions for confounding of interactions in logistic regression. Our main result is that failure to adjust for a risk factor U will confound the $E1 * E2$ interaction estimates if and only if (C) the association between $E1$ and U depends on the value of $E2$, which implies an interaction between the two exposures for U. Interestingly, our conditions A and C extend the classic conditions, according to which a confounder of the main effect of a single exposure has to be associated with both, respectively, the outcome and the exposure,[26,36] to the context of an interaction between two exposures. Indeed, our condition C implies that U has to be associated with the interaction term $E1 * E2$, which, in our context, represents the "exposure" of interest.

These findings should be considered when interpreting the results of observational studies of interactions in terms of potential bias due to unmeasured confounding. The important consequence is that a confounder for the marginal effect of $E1$ and/or $E2$, which, however, does not meet our condition C, is *not* a confounder for the interaction between the two exposures. Accordingly, researchers concerned about potential confounding of the estimate of the interaction of interest should focus on identifying, based on the literature and/or substantive considerations, those risk factors for Y, for which their association with $E1$ is expected to vary depending on E2. Once identified, such variables should be ideally adjusted for in the regression model used to estimate the interaction of interest. If such confounders (U) are not available, then our simulation results, summarized in Tables 3–5, will help predict if and how the resulting estimates may be biased depending on the expected associations of U with Y and the $E1 * E2$ interaction. Consider, for example, education (U) in a study of the interaction between sex ($E1$) and age ($E2$). Among younger subjects ($E2 = 0$), women ($E1 = 1$) may have higher average education than men ($E1 = 0$), but among older subjects ($E2 = 1$), men may have higher education than women. This implies a strong interaction $E1 * E2$ for U, but the lack of clear marginal differences in U between levels of $E1$ or $E2$ (e.g., sex-related differences in younger versus older subjects will approximately cancel out). If so, for any outcome (Y) associated with education, a failure to adjust for education may bias the estimated effect of sex * age interaction, even if U will not be considered a confounder of the main effect of sex on Y.

Our analytical results were confirmed by the simulation experiments and illustrated by real-life analyses of a case–control study of lung cancer. Simulations indicate also that the bias due to a failure to adjust for U that meets our conditions A and C, increases with an increase in the following: (i) prevalence of the confounder; (ii) strength of its association with the outcome; and (iii) strength of the interaction between the two exposures for U, that is, differences between $E1$–U associations across the strata of E2. Furthermore, omitting U, which meets the aforementioned conditions, may induce a serious risk of type I error, that is, detecting a spurious interaction when none truly exists. Additional simulations demonstrate that our conditions apply regardless of whether exposures and/or unobserved confounder represents binary or continuous variables. Both simulation results and empirical lung cancer analyses suggest that the impact of a failure to adjust for an unmeasured risk factor depends mostly on whether and to what extent it violates our condition C; that is, this impact reflects mostly unmeasured confounding rather than noncollapsibility of odds ratios.

Knol *et al*. and Kaufman pointed out several limitations of the current practice of interaction assessment[7,13] but did not mention unmeasured confounding. Among 30 articles we reviewed that investigated potential interactions, the most common approach was to first build multivariable regression models to provide valid estimates for the main effect(s) of the exposure(s) in the interaction term on the outcome and then add the hypothesized interaction product term(s) to the model. Thus, covariates were typically selected based on concerns about potential confounding of the main effects of the exposures. The same practice permeates recent pharmaco-epidemiological studies of drug–drug or drug–covariate interactions.[2,3] Yet, our simulations demonstrate that a variable that does *not* meet the classical conditions for a confounder of the main effects of either exposure may still act as an important confounder of their interaction. Thus, the current practice might lead to omission of important confounders of the interaction, especially when parsimonious models are preferred.

We considered statistical interaction only on a multiplicative scale, arguably the most common form of interaction investigated in current epidemiological practice. Thus, our results pertain to logistic regression-based OR and RR estimates. Whereas our results cannot be automatically extrapolated to interactions on the additive scale,[7,10] we expect that our analytical framework and simulation methods could be also useful in such investigations.

There is no consensus regarding how to estimate and report interaction effects.[13] Given the growing importance of interaction assessments, we hope that our study will stimulate further methodological investigation of a topic that has been rather neglected, and a more informed interpretation of empirical results of interaction analyses.

DOI: 10.1002/pds

ETHICS STATEMENT

The Montreal study was approved by the ethics committees of the Institut national de recherche scientifique-IAF, of McGill University, and of each hospital in the area that provided patients. Further, each participant provided informed consent.

ACKNOWLEDGEMENTS

Source of financial support: This research was supported by the Canadian Institutes for Health Research (CIHR) grant #MOP-81275 (PI: M. Abrahamowicz), by CIHR grant #MOP-14704 (PI: J Siemiatycki), and by the Guzzo-Cancer Research Society Chair in Environment and Cancer (chairholder: J Siemiatycki).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

KEY POINTS

- Failure to adjust for a risk factor *U* results in a biased estimate of the interaction between exposures *E*1 and *E*2 on a binary outcome if the association between *U* and *E*1 varies depending on the value of *E*2, or *vice versa*.
- The resulting bias increases with increasing strength of the interaction through which *E*2 modifies the association between *U* and *E*1.
- A variable *U* may act as an important confounder of the *E*1–*E*2 interaction even if *U* does not meet the classical conditions for being a confounder for the main effects of either *E*1 or *E*2.
- Epidemiological studies of interactions should attempt to identify—as potential confounders—those risk factors whose association with one of the "exposures" is likely to be modified by the other "exposure."

REFERENCES

1. Pottegard A, de Pont Christensen R, Wang SV, Gagne JJ, Larsen TB, Hallas J. Pharmacoepidemiological assessment of drug interactions with vitamin K antagonists. *Pharmacoepidemiol Drug Saf* 2014; **23**(11): 1160–7.
2. Yue Z, Shi J, Jiang P, Sun H. Acute kidney injury during concomitant use of valacyclovir and loxoprofen: detecting drug–drug interactions in a spontaneous reporting system. *Pharmacoepidemiol Drug Saf* 2014; **23**(11): 1154–9.
3. Rowan CG, Brunelli SM, Munson J, *et al.* Clinical importance of the drug interaction between statins and CYP3A4 inhibitors: a retrospective cohort study in The Health Improvement Network. *Pharmacoepidemiol Drug Saf* 2012; **21**(5): 494–506.
4. Stevens V, Dumyati G, Brown J, Wijngaarden E. Differential risk of *Clostridium difficile* infection with proton pump inhibitor use by level of antibiotic exposure. *Pharmacoepidemiol Drug Saf* 2011; **20**(10): 1035–42.
5. Behr S, Andersohn F, Garbe E. Risk of intracerebral hemorrhage associated with phenprocoumon exposure: a nested case–control study in a large population-based German database. *Pharmacoepidemiol Drug Saf* 2010; **19**(7): 722–30.
6. Graham DJ, Zhou EH, McKean S, *et al.* Cardiovascular and mortality risk in elderly Medicare beneficiaries treated with olmesartan versus other angiotensin receptor blockers. *Pharmacoepidemiol Drug Saf* 2014; **23**(4): 331–9.
7. Kaufman JS. Interaction reaction. *Epidemiology* 2009; **20**(2): 159–60.
8. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 2004; **57**(3): 229–36.
9. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med* 2002; **21**(19): 2917–30.
10. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol* 1980; **112**(4): 467–70.
11. Saracci R. Interaction and synergism. *Am J Epidemiol* 1980; **112**(4): 465–6.
12. Abrahamowicz M, Beauchamp ME, Fournier P, Dumont A. Evidence of subgroup-specific treatment effect in the absence of an overall effect: is there really a contradiction? *Pharmacoepidemiol Drug Saf* 2013; **22**(11): 1178–88.
13. Knol MJ, Egger M, Scott P, Geerlings MI, Vandenbroucke JP. When one depends on the other: reporting of interaction in case–control and cohort studies. *Epidemiology* 2009; **20**(2): 161–6.
14. Greenland S. Interactions in epidemiology: relevance, identification, and estimation. *Epidemiology* 2009; **20**(1): 14–7.
15. Walker AM. Confounding by indication. *Epidemiology* 1996; **7**(4): 335–6.
16. Rothman KJ, Greenland S, Last TL. Measures of effect and measures of association. In *Modern Epidemiology* (3rd edn). Lippincott-Williams-Wilkins: Philadelphia, 2008; 58–9.
17. Hlatky MA, Boothroyd DB, Bravata DM, *et al.* Coronary artery bypass surgery compared with percutaneous coronary interventions for multivessel disease: a collaborative analysis of individual patient data from ten randomised trials. *Lancet* 2009; **373**(9670): 1190–7.
18. Cote ML, Chen W, Smith DW, *et al.* Meta- and pooled analysis of GSTP1 polymorphism and lung cancer: a HuGE-GSEC review. *Am J Epidemiol* 2009; **169**(7): 802–14.
19. Sharma AJ, Cogswell ME, Li R. Dose–response associations between maternal smoking during pregnancy and subsequent childhood obesity: effect modification by maternal race/ethnicity in a low-income US cohort. *Am J Epidemiol* 2008; **168**(9): 995–1007.
20. Tobin MD, Kahonen M, Braund P, *et al.* Gender and effects of a common genetic variant in the NOS1 regulator NOS1AP on cardiac repolarization in 3761 individuals from two independent populations. *Int J Epidemiol* 2008; **37**(5): 1132–41.
21. Onufrak SJ, Bellasi A, Cardarelli F, *et al.* Investigation of gender heterogeneity in the associations of serum phosphorus with incident coronary artery disease and all-cause mortality. *Am J Epidemiol* 2009; **169**(1): 67–77.
22. Kosiborod M, Inzucchi SE, Goyal A, *et al.* Relationship between spontaneous and iatrogenic hypoglycemia and mortality in patients hospitalized with acute myocardial infarction. *JAMA* 2009; **301**(15): 1556–64.
23. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci* 1999; **14**(1): 29–46.
24. Samuels ML. Matching and design efficiency in epidemiological studies. *Biometrika* 1981; **68**(3): 577–88.
25. Gail MH, Wieand S, Piantadosi S. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika* 1984; **71**(3): 431–44.
26. Rothman KJ. Biases in study design. In *Epidemiology: An Introduction*. Oxford University Press: New York, 2002; 108.
27. Szklo M, Nieto F. Identifying noncausal associations: confounding. In *Epidemiology: Beyond the Basics* (2nd edn). Mass, Jones & Bartlet Publishers: Boston, 2007; 151–81.
28. Egberts AC, Meyboom RH, van Puijenbroek EP. Use of measures of disproportionality in pharmacovigilance: three Dutch examples. *Drug Saf* 2002; **25**(6): 453–8.
29. Young J, Xiao Y, Moodie EE, *et al.* Effect of cumulating exposure to abacavir on the risk of cardiovascular disease events in patients from the Swiss HIV cohort study. *J Acquir Immune Defic Syndr* 2015; **69**(4): 413–21.
30. Miettinen OS, Cook EF. Confounding: essence and detection. *Am J Epidemiol* 1981; **114**(4): 593–603.
31. Hernan MA, Clayton D, Keiding N. The Simpson's paradox unraveled. *Int J Epidemiol* 2011; **40**(3): 780–5.
32. Gail M. Adjusting for covariates that have the same distribution in exposed and unexposed cohorts. In *Modern Statistical Methods in Chronic Disease*

Epidemiology, Moolgavkar SH, Prentice R L (eds.), Wiley: New York; 1986; pp. 3–18.
33. Ramanakumar AV, Parent ME, Siemiatycki J. Risk of lung cancer from residential heating and cooking fuels in Montreal, Canada. *Am J Epidemiol* 2007; **165**(6): 634–42.
34. Benedetti A, Parent M, Siemiatycki J. Consumption of alcoholic beverages and risk of lung cancer: results from two case–control studies in Montreal, Canada. *Cancer Causes Control* 2006; **17**: 469–80.
35. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. *Stat Med* 2005; **24**(11): 1713–23.
36. Gordis L. *Epidemiology* (4th edn). Elsevier/Saunders: Philadelphia, 2009; xv pp. 375.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site.

DOI: 10.1002/pds