# Confounding time dependent or not: introspection

Ranadip Chowdhury[1], Kaushik Sarkar[2]

[1]Young Investigator, Centre for Health Research and Development (CHRD), Society for Applied Studies (SAS), New Delhi, India.
[2]Department of Community Medicine, All India Institute of Hygiene and Public Health, Kolkata, West Bengal, India.
Correspondence to: Ranadip Chowdhury, E-mail: ranadip84@gmail.com

Received December 21, 2015. Accepted February 3, 2016

**Abstract**

Longitudinal studies, where data are repeatedly collected on one subject over a period, are common in medical research. When effect of a time-varying exposure on an outcome of interest is measured at different time points, standard statistical methods fail to give robust estimate in the presence of time-dependent confounders. There is alternative method avoid, that is, inverse probability weighted estimation of marginal structural models the problems associated with standard approaches.

**KEY WORDS:** Time-dependent confounding, marginal structural models, bias

## Introduction

Estimated effects of an exposure on health outcome vary between studies. For example, association between breast-feeding and childhood asthma in observational studies has shown both harmful and beneficial effects.[1–5] Similarly, effect of breast-feeding on atopic dermatitis is inconsistent from observational studies.[6,7] Reverse causality, selection biases, and unmeasured confounding may be the reasons for these inconsistent results.[8,9] To measure the effect of breast-feeding on child health outcomes from observational studies, a range of confounders, that is, maternal, pregnancy, and perinatal risk factors for the outcome of interest are generally adjusted.[2–5] Time-dependent confounding could be another reason for these inconsistent findings from observational studies. Time-dependent confounder is a variable that is associated with current exposure and future outcome, predicted by previous exposure, and predicts current exposure.[10] Here, conventional statistical methods produce biased estimates of exposure effect, because they fail to account for the time-dependent nature of the confounders and exposures.[11]

Conventional statistics are readily implemented provided that confounders are measured once, at the start of the follow-up. The relationship between breast-feeding and subsequent child health outcomes may also be subject to time-dependent confounding, when risk factors for the outcome both predict and are changed by breast-feeding. For example, infant weight may influence the probability that a mother continues breast-feeding. In turn, breast-feeding (compared with formula feeding) may affect weight gain.[12]

Time-varying effects of exposure have been considered in the epidemiologic and statistical literature.[13–15] The aim of the review is to extend the literature with the understanding of time-varying third variable model by elucidating the concept of time-dependent confounding variable and how to adjust those variables to infer the association between exposure and outcome.

| Access this article online | |
|---|---|
| Website: http://www.ijmsph.com | Quick Response Code: |
| DOI: 10.5455/ijmsph.2016.21122015366 | |

## Confounding

Confounding refers to a situation where exposure and outcome share a common path, which may be represented by one variable or multiple variables. Let A (0) represent exposure (breast-feeding) at time 0 (using values in parentheses to represent time), B (2) outcome (allergy) at time 2, and C (0) a confounding variable (infant weight) occurring temporally before A (0) that has a direct causal effect on both A (0) and B (2); note that, in Figure 1a, A (0) has no causal effect on B (2). We need to consider substantive knowledge in decisions on adjustment for confounders. In Figure 1a, control for C (0) through regression, stratification, or restriction provides a consistent estimate of the C (0)—conditional causal effect of exposure A (0) on outcome B (2). A series of extensions of the confounding definition to settings involving time-varying variables and effects will be considered in the following sections.

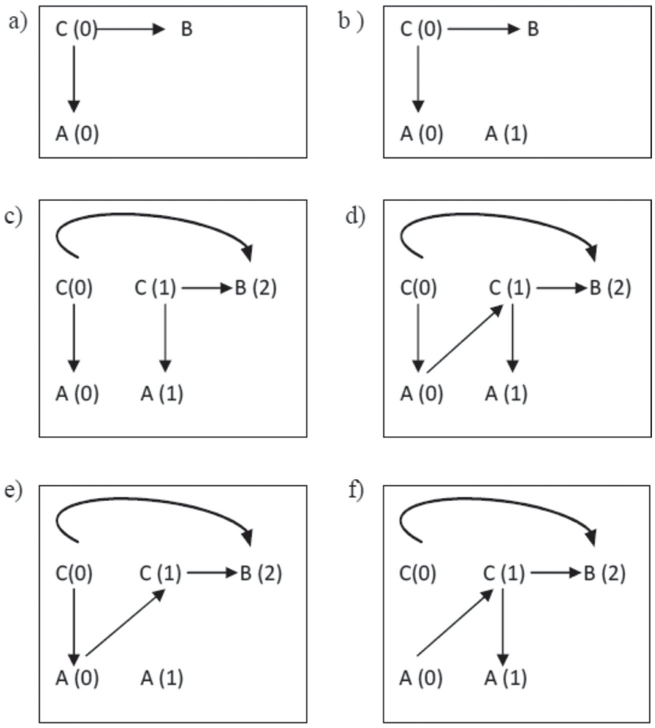


**Figure 1:** Causal diagrams representing confounding (a), time-modified confounding by a time-fixed covariate (b), time-varying confounding (c), time-varying confounding affected by prior exposure (d), and time-modified confounding (e and f)

### Time-Modified Confounding

Here exposure A now varies over time, but C (0) confounds only the A (0)–B (2) relationship. This is a case of time-modified confounding. Control for C (0) will give a consistent estimate of the C (0)—conditional causal effect of A (0) on B (2), while the effect of A (1) on B (2) can be estimated consistently from the crude (unadjusted) model [Figure 1b].

### Time-Varying Confounding

Figure 1c shows A (0) and A (1) representing exposure at times 0 and 1, and C (0) and C (1) represent time-varying confounding variables measured temporally before times $t = 0$ and $t = 1$, respectively. At each time point, the confounders have a direct causal effect on exposure. Now to estimate the total (i.e., direct and indirect) causal effect of A (0) on B (2), adjustment for C (0) is necessary. To estimate the total causal effect of A (1) on B (2), we must adjust for C (1). In this situation, standard statistical methods (e.g., Cox regression with time-varying exposure and covariates) can consistently estimate, as previously defined, the C (*t*)—conditional causal effect of exposure A (*t*), $t = 0$ or 1, on outcome B (2).

Figure 1d describes an extension of Figure 1c, where now the time-varying confounder is affected by prior exposure. In this case, adjustment for C (1) is necessary to estimate the effect of A (1) but blocks a causal effect of A (0). Estimation of the total effect of A (*t*) will be biased when standard statistical methods will be used. Robins et al.[16] have proposed a series of methods to address this problem, including marginal structural models (MSM).

### Time-Modified Confounding by Time-Varying Factors

Figure 1e describes a further extension of Figure 1c, which illustrates time-modified confounding. In this case, the effect of the confounding variables C (*t*) on exposure A (*t*) differs over time *t*; in the case of Figure 1e, there is a direct causal effect of C (0) on A (0), but there is no direct causal effect of C (1) on A (1). Figure 1f describes a companion scenario of time-modified confounding, where C (0) has no direct causal effect on A (0), although C (1) does have a direct causal effect on A (1).

In both the examples (i.e., Figure 1e and 1f), the causal effect of A (*t*) on B (2) could be consistently estimated by standard methods (such as linear or logistic regression). For example, in Figure 1e, adjusting for C (0) is sufficient to control confounding and to provide an unbiased estimate of the causal effect. In Figure 1f, the simple cross-tabulation of A (0) and B (2) would provide an unbiased estimate of the causal effect of A (*t*) on B (2). However, these approaches are based on the knowledge that the hypothesized diagram is correct. In practice, one may be unlikely to estimate the effect of A (*t*) without using all measured exposures and confounding information. Such adjustment may introduce bias. In practice, these simple solutions would fail if there was a causal effect of C (1) on A (1).[16] MSM can provide consistent estimates in either case.

## Methods

MSM aim to estimate the effect of exposure on outcome by appropriate control for the effects of time-dependent confounders. The model is fitted in a two-stage process in which

1. each subject's probability of having their own exposure history is used to derive inverse-probability-of-exposure weights (IPTW), and
2. the exposure–outcome association is estimated in a regression model that is weighted using the IPTWs.

MSM technique creates at each point of time a pseudo-population of counterfactuals (a hypothetical population in which all patients seem as exposed and unexposed to an exposure), in which, time-invariant and time-dependent confounders are balanced, and therefore, causal association between the exposure and the outcome is the same as in the original study population.[17] The pseudo-population is created by weighting every patient in the population by the inverse of the conditional probability of being exposed to the exposure that the study participant actually exposed. The technique compares two counterfactuals: outcome of the entire study population exposed to the exposure and outcome of the entire study population not exposed to the exposure. Thus, it gives valid causal interpretations between the exposure and the outcome.

**Strengths[18]**

1. Re-weighting to achieve a pseudo-sample unaffected by confounding is intuitive and relatively easy to explain.
2. The analysis is easily implementable using weighted versions of standard routines.
3. MSM cope well with different sorts of outcome variable. It is possible to fit logistic MSM for binary outcomes and Cox MSM for time-to-event outcomes.

**Limitations[18]**

1. Inverse weighting can be unstable and inefficient if there are extreme weights.
2. Continuous exposures are difficult to handle.
3. Possible interactions between exposure and time-varying covariates cannot be explored because the MSM are marginal with respect to the latter.

## Discussion

In this technical note, we have highlighted the issues associated with time-dependent confounding. We have described the application of MSM to combat time-dependent confounding variable. The impact of time-dependent confounding is an important issue that must be seriously considered by researchers analyzing data from longitudinal observational studies with time-varying exposures. Appropriate methods should be used to evaluate the impact of potential time-dependent confounding. Even if adjustment for a large number of baseline covariates does not affect estimated exposure effects, there may still be situations in which time-dependent confounding will cause biased results. Ideally, before conducting a study on time-dependent exposures, researchers should consider the potential for time-dependent confounding, identify the (most important) time-dependent confounders, and measure those repeatedly.

## References


1. Kramer MS, Matush L, Vanilovich I, Platt R, Bogdanovich N, Sevkovskaya Z, et al; PROBIT Study Group. Effect of prolonged and exclusive breast feeding on risk of allergy and asthma: cluster randomized trial. BMJ 2007;335(7624):815.
2. Kull I, Melen E, Alm J, Hallberg J, Svartengren M, van Hage M, et al. Breast-feeding in relation to asthma, lung function, and sensitization in young schoolchildren. J Allergy Clin Immunol 2010;125(5):1013–9.
3. Scholtens S, Wijga AH, Brunekreef B, Kerkhof M, Hoekstra MO, Gerritsen J, et al. Breast feeding, parental allergy and asthma in children followed for 8 years. The PIAMA birth cohort study. Thorax 2009;64(7):604–9.
4. Sears MR, Greene JM, Willan AR, Taylor DR, Flannery EM, Cowan JO, et al. Long-term relation between breastfeeding and development of atopy and asthma in children and young adults: a longitudinal study. Lancet 2002;360(9337):901–7.
5. Giwercman C, Halkjaer LB, Jensen SM, Bønnelykke K, Lauritzen L, Bisgaard H. Increased risk of eczema but reduced risk of early wheezy disorder from exclusive breast-feeding in high-risk infants. J Allergy Clin Immunol 2010;125(4):866–71.
6. Kramer MS, Chalmers B, Hodnett ED, Sevkovskaya Z, Dzikovich I, Shapiro S, et al; PROBIT Study Group (Promotion of Breastfeeding Intervention Trial). Promotion of breastfeeding intervention trial (PROBIT): a randomized trial in the Republic of Belarus. JAMA 2001;285(4):413–20.
7. Yang YW, Tsai CL, Lu CY. Exclusive breastfeeding and incident atopic dermatitis in childhood: a systematic review and meta-analysis of prospective cohort studies. Br J Dermatol 2009;161(2):373–83.
8. Kramer MS, Guo T, Platt RW, Shapiro S, Collet JP, Chalmers B, et al; PROBIT Study Group. Breastfeeding and infant growth: biology or bias? Pediatrics 2002;110(2 Pt 1):343–7.
9. Kramer MS, Moodie EE, Dahhou M, Platt RW. Breastfeeding and infant size: evidence of reverse causality. Am J Epidemiol 2011;173(9):978–83.
10. Ali AK. Methodological challenges in observational research: a pharmacoepidemiological perspective. Br J Pharmaceutical Res 2013;3(2):161–75.
11. Suarez D, Borràs R, Basagaña X. Differences between marginal structural models and conventional models in their exposure effect estimates. A systematic review. Epidemiol 2011;22(4): 586–8.
12. Moodie EEM, Platt RW, Kramer MS. Estimating response-maximized decision rules with applications to breastfeeding. J Am Stat Assoc 2009;104(485):155–65.
13. Platt RW, Joseph KS, Ananth CV. A proportional hazards model with time-dependent covariates and time-varying effects for analysis of fetal and infant death. Am J Epidemiol 2004;160(3): 199–206.
14. Abrahamowicz M, MacKenzie T, Esdaile JM. Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. J Am Stat Assoc 1996;91(436):1432–9.
15. Altman DG, De Stavola BL. Practical problems in fitting a proportional hazards model to data with updated measurements of the covariates. Stat Med 1994;13(4):301–41.
16. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550–60.
17. Hernán MA, Brumback B, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. Stat Med 2002;21(12):1689–709.
18. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JAC. Methods for dealing with time-dependent confounding. Stat Med 2013;32(9):1584–618.


**How to cite this article:** Chowdhury R, Sarkar K. Confounding time dependent or not: introspection. Int J Med Sci Public Health 2016;5:1973-1975

**Source of Support:** Nil, Conflict of Interest: None declared.