Nephrol Dial Transplant (2017) 32: ii84–ii90
doi: 10.1093/ndt/gfw341
Advance Access publication 11 February 2017

ndt
Nephrology Dialysis Transplantation

Full Review

Marginal structural models in clinical research: when and how to use them?

Tyler Williamson[1,2] and Pietro Ravani[1,3]

[1]O'Brien Institute of Public Health, Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada, [2]Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada and [3]Libin Cardiovascular Institute of Alberta, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

Correspondence and offprint requests to: Tyler Williamson; E-mail: tyler.williamson@ucalgary.ca, Pietro Ravani; E-mail: pravani@ucalgary.ca

ABSTRACT

Marginal structural models are a multi-step estimation procedure designed to control for the effect of confounding variables that change over time, and are affected by previous treatment. When a time-varying confounder is affected by prior treatment standard methods for confounding control are inappropriate, because over time the covariate plays both the role of confounder and mediator of the effect of treatment on outcome. Marginal structural models first calculate a weight to assign to each observation. These weights reflect the extent to which observations with certain characteristics (covariate values) are under-represented or over-represented in the sample with the respect to a target population in which these characteristics are balanced across treatment groups. Then, marginal structural models estimate the outcome of interest taking into account these weights. Marginal structural models are a powerful method for confounding control in longitudinal study designs that collect time-varying information on exposure, outcome and other covariates.

Keywords: bias, confounding, inverse probability of treatment weight, longitudinal study design, marginal methods

INTRODUCTION

The relationship between treatment and outcome is often affected by other factors. For example, if treatment is use of beta-blockers and outcome is mortality post-myocardial infarction, and the treated group is older than the untreated group, age may confound the relationship between treatment and outcome by making benefits from beta-blockers look smaller [1]. When a treatment is not assigned randomly, as is an intervention in a randomized controlled trial, or more generally in observational studies of an exposure and an outcome, confounding needs to be accounted for in order to obtain an unbiased estimate of the relationship of interest. Confounding occurs whenever a factor is associated with (or causes) both treatment and outcome, and is not on the causal path between treatment and outcome. In the example above, age is a confounder because it is associated with (affects as opposed to being affected by) the use of beta-blockers, and is not on the causal pathway between use of beta-blockers and death (Figure 1).

There are many ways to correctly disentangle the effect of the treatment from the effect of the confounder, including study design strategies (e.g. randomized controlled trials) and analytic strategies (e.g. stratification, regression or matching methods). Randomized controlled trials are the best way to control for confounding because they balance both known and unknown (unmeasured) confounding factors. However, randomized controlled trials are not suitable for all research questions, including post-marketing studies of interventions, or studies of non-modifiable risk factors (i.e. ethnicity, age or sex). Given the popularity of observational studies, confounding has been given considerable attention in the biostatistical and epidemiological literature. For example, use of catheters for hemodialysis access is associated with increased mortality as compared with use of fistulas [2]. In one study the risk for death in catheter users as compared with fistula users was lower in analyses adjusted for

several comorbid factors that are also more common in catheter users (i.e. increase the likelihood of using catheters) than in fistula users as compared with analyses including fewer or no comorbid factors [3]. In this study the authors used regression techniques to remove the effect of confounding on the relationship of interest [4, 5]. Other methods can be used to increase group comparability, for example, by stratification, restriction or propensity score matching [6].

In some situations, however, other variables may have different effects on either the intervention (or the exposure) and the outcome, or both. In the example above, change in sympathetic nervous activity may mediate the effect of beta-blockers on mortality (Figure 1). If the reduced activity of the sympathetic nervous system mediates fully or partly the effect of beta-blockers on mortality, an analysis adjusted for markers of sympathetic nervous activity (e.g. heart rate or blood pressure) would remove such indirect effects. Measures of sympathetic nervous activity would not be included in standard outcome analysis, unless the objective was to study mediation effects [7].

Other variables may affect the relationship between treatment (exposure) and outcome. Figure 2 summarizes the complex relationship between maternal deprivation and low birth weight and the effects of other variables: confounding by maternal age; mediation by diet (similar to the sympathetic nervous activity control in the previous example); modification or interaction by smoking; and independent effects of mother height. An analysis that fails to properly handle these different effects may provide a biased estimate of the true effect of the intervention (exposure) on outcome. One popular example is the low-

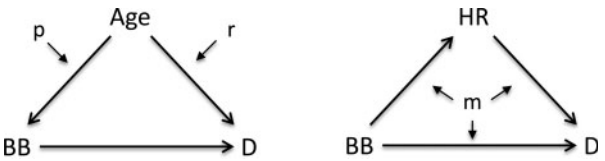


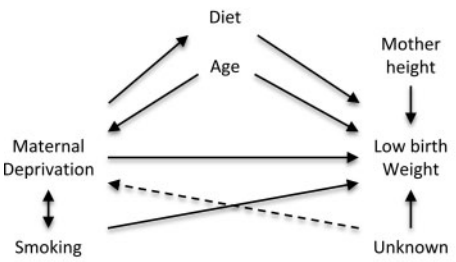
FIGURE 1: Confounders and mediators. A confounding variable is a third variable affecting both treatment (beta-blockers; BB) and outcome (death; D), without being on the pathway between treatment and outcome. Standard confounding adjustment consists of removing the influence of confounders on the relationship of interest in order to isolate true effects by means of propensity score methods (p) or regression analysis (r), for example. A mediator is a third variable that mediates (partially or fully) the relationship between treatment and outcome [e.g. heart rate (HR) as a measure of sympathetic nervous system activity]. Mediation analysis (m) is used to estimate the components of the total effect of treatment on outcome.

weight paradox, which refers to the apparently paradoxical observation of lower infant mortality in children born to tobacco smoking mothers as compared with non-smoking mothers (Figure 3). Accounting for (i.e. conditioning on) birth weight introduces a bias known as collider stratification bias, which in this case changes the sign of the effect of smoking on mortality. In this example, birth weight is a collider because it is affected by smoking and genetic defects (two effects colliding on the same variable).

CONFOUNDING IN LONGITUDINAL DESIGNS

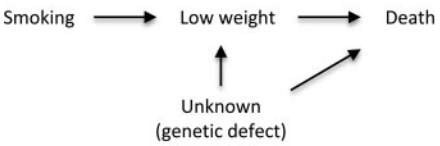
In longitudinal studies with repeated measures of confounding factors, mediators or colliders, these complex relationships are commonly encountered (Figure 4). Over time the same variable may be a confounder, a collider or a mediator (i.e. time-varying confounder affected by previous treatment), and standard methods to address confounding will produce biased results. For example, the estimate of the association between use of beta-blockers and mortality adjusted for baseline smoking or age but without control of baseline or time-varying blood pressure values will produce biased estimates because blood pressure values are associated with the outcome, and may affect the use of beta-blockers both at baseline and over time. Control for the baseline values of blood pressure only will give biased estimates because a confounding effect occurring over time is ignored. On the other hand, standard methods to control for time-updated measurements of blood pressure affected by prior use of beta-blockers will give biased estimates by removing (adjusting away) the effects of beta-blockers mediated through blood pressure control, or by inducing selection bias by conditioning on a collider (collider-stratification bias).

Marginal structural models aim to appropriately control for the effects of time-varying confounders that are affected by prior treatment (exposure). Marginal structural models are called *marginal* because they use the marginal distribution of the treatment (or exposure) variable at any time point to weight its relationship with outcome. Marginal in statistics means 'unconditional'—i.e. not conditional on other variables, like the probabilities of one variable on the margins of a contingency table of two discrete variables (see below). These models are called *structural* because models exploring causal relationships are referred to as structural in the econometric and social sciences literature [8].



Relations with low birth weight (outcome)	
Variable	*Type of relations*
Maternal deprivation	Exposure
Smoking	Moderator
Age	Confounder
Diet	Mediator
Mother height	Independent variable
Unknown	Potential confounder

FIGURE 2: Multiple factors affecting the relationship between exposure and outcome.



Collider stratification bias example		
	Smoking	*No smoking*
Low weight	63/900 (0.07)	8/100 (0.08)
No low weight	4/100 (0.04)	45/900 (0.05)
Totals	67/1000 (0.067)	53/1000 (0.053)

FIGURE 3: Collider-stratification bias.

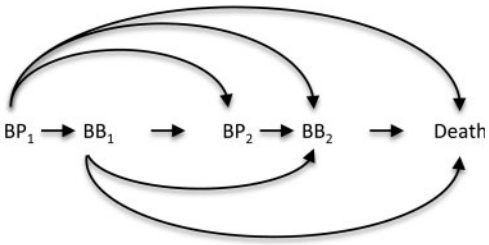


FIGURE 4: Time-varying confounding. In some longitudinal studies variables are measured more than once over time. A third variable (e.g. blood pressure; BP) in these situations is associated with (i) previous treatment (beta-blockers; BB), (ii) subsequent treatment (BB) and (iii) outcome.

FULL REVIEW

MARGINAL STRUCTURAL MODELS

Causal inference

Marginal structural models are a class of statistical models used for causal inference in epidemiology. Although much of the literature on marginal structural models focuses on counterfactuals and causal effects, using counterfactuals to explain marginal structural models can be misleading. Ideally, the true causal effect of a treatment (or exposure) would be estimated by comparing what actually happened to a subject given a certain treatment with what would have happened to the same subject if (contrary to fact) they had received the alternative treatment. Since we can rarely observe the same patients under both conditions (even in a crossover randomized controlled trial carry-over effects induce systematic differences), the only way to determine causal effects is to compare two groups that are identical in every way except that one group was treated and the other was not. Marginal structural models (as other causal models) attempt to fully adjust for measured confounders to enhance group comparability and estimate causal effects in a similar way [8]. While these methods have been found to replicate results of randomized controlled trials when all their assumptions are met [1], only randomization can ensure comparability of treatment groups by balancing both measured and unmeasured confounding factors. Examples in nephrology include studies of the relationship between treatment with erythropoietin and mortality accounting for the time-varying confounding effect of anemia [9], or studies of the effect of dialysis dose on mortality accounting for the time-varying confounding effect of laboratory markers of uremia [10].

Marginal methods to control confounding

Methods for dealing with confounding can be grouped into two major approaches, the conditional approach and the marginal approach. Both approaches attempt to estimate the effect of a treatment (exposure) on an outcome by balancing the distribution of confounders across level of treatment (exposure) to reflect their overall distributions in the target population (of which the study population is assumed to be an unbiased sample).

The conditional approach estimates the effect on outcome for each level of treatment (exposure) as a weighted average of the effects over levels of confounders (conditional on confounders). This is obtained by either stratification or modeling. Stratification assumes no residual confounding within stratum. However, the possibility to increase the number of covariate combinations is limited, as small sample size for some combinations of confounders would affect the estimation process. Modeling requires correct model specification, including shape of the relationship of each model variable with the outcome (e.g. linear, log-linear, exponential, etc.—how the outcome changes as the independent variables varies), and consideration of effect modifications (interactions) [4, 5]. The marginal approach requires balancing the confounders across levels of treatment (exposure) before outcome modeling. In marginal structural models this balancing is accomplished by assigning weight to each observation to obtain a balanced (re-weighted) sample, as opposed to conditioning on the characteristics (covariates) of each observation. The estimated effects of an intervention (or exposure) on outcome obtained using this approach are called 'marginal' (population average) effects, as opposed to conditional (individual) effects.

Revisiting the example of the relationship between beta-blockers and mortality, the conditional approach estimates the risk of death by use of beta-blockers, and averages these estimates across levels of (conditional on) blood pressure in the whole sample. The marginal approach seeks to first balance the distribution of blood pressure with respect to the use of beta-blockers, and then compares mortality between treatment groups. Weights are calculated first to reflect the extent to which each observation is under-represented or over-represented in the original sample with respect to a target population in which potential confounders are balanced across treatment groups. To increase group comparability, in outcome analysis a re-weighted sample is obtained in which the weight of observations that are under-represented in the original sample is increased,

and the weight of those over-represented is reduced. Sample re-weighting removes the imbalances induced by uneven distribution of confounding across treatment (exposure) groups [6].

ASSUMPTIONS OF MARGINAL STRUCTURAL MODELS

The main assumptions of marginal structural models are exchangeability, consistency, positivity and correctness of the weight-generating model [8]. *Exchangeability* implies absence of unmeasured confounding inducing correlation between treatment (exposure) and residuals. While this is a requirement of any statistical model, it cannot be tested in any setting using observational data [11]. If this assumption is met (i.e. using randomized controlled designs), causal inference is possible from marginal structural models. *Consistency* requires that the outcome observed for each individual is precisely the causal outcome under their observed treatment history, which is also difficult to verify. This assumption would be violated in the presence of misclassification bias. *Positivity* requires that the probability of treatment is neither zero nor one for each combination of covariates. That is, the distribution of treatment must vary across every unique covariate combination (i.e. the confounders cannot determine the treatment or non-treatment status perfectly). Situations where the treatment assignment is deterministic (for each covariate combination treatment is either zero or one), or nearly deterministic (close to zero or one), can dramatically influence the estimation of the model and should be carefully examined. This can happen, for example, if there are levels of confounders where individuals could not possibly be treated, such as the time period before a particular treatment existed, recommendations from guidelines or established contraindications for treatment (structural zeroes). Violation of this assumption leads to biased estimates. A related problem is the occurrence of random zeroes, which are zero probabilities resulting by chance, usually due to small sample sizes in some covariates levels. While additional categories of confounders may provide better adjustment for confounding, the resulting increase in random zeroes can increase the bias and variance of the estimate effect. Sensitivity analysis can be used to examine the bias-variance trade off of combining small groups or excluding weak confounders from the models. The final assumption of marginal structural models is that the model used to estimate the weights must be correctly specified (and its assumptions verified and met), which has similar roots in essentially all statistical models [12]. The importance of other assumptions is debated [13].

ESTIMATION IN MARGINAL STRUCTURAL MODELS

Of the different available methods that have been proposed to address time-varying confounding with marginal structural models [14], inverse-probability-of-treatment weighting (IPTW) is the most commonly used [15]. Marginal structural model estimation via IPTW occurs in two stages (Figure 5). As

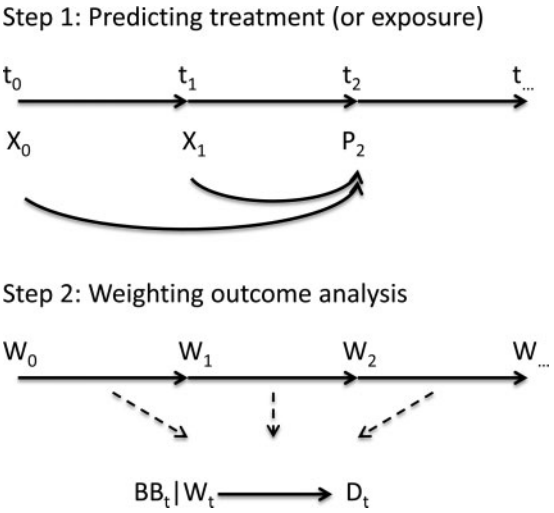


FIGURE 5: Analysis steps in marginal structural models. Step 1: at each time point the probability (P) of using beta-blockers (BB) is estimated based on previous values of the covariates (X), including potential time-varying confounders. Step 2: the probabilities obtained in the previous analysis step are used to generate weights (W) to address confounding by weighting the influence of each observation in the estimation of the effects.

a preliminary step, researchers choose the time scale, and define time, event, treatment and other covariates, and their relationships (Figures 2 and 4). In the first stage IPTW are calculated; in the second stage the outcome model is fit, including sensitivity analyses that take into account weight distributions.

Inverse-probability-of-treatment weighting

Treatment weights. Treatment weights are calculated based on the inverse of each individual's probability of the treatment they were actually receiving at each observation time point, given their covariate history. These inverse probability weights are then stabilized using probabilities of the treatment history given their baseline information only to reduce their variability (see Supplementary data, Appendix for details).

Censoring weights. In longitudinal studies some participants are lost to follow-up throughout the study period for any number of reasons including factors related to the treatment, or a competing event. This attrition will influence the results if there is a systematic difference between those that are lost to follow-up and those that remain in the study across treatment (exposure) groups, and the reason for leaving the study is related to the outcome. The mechanism that causes this informative censoring can be thought of as another form of a time-dependent confounder because the probability of censoring depends on the outcome the subject would have had in the absence of censoring. In marginal structural models this form of bias is dealt with by including censoring weights in the estimation process (Supplementary data, Appendix), which are used to recreate the population that would be seen without dropouts (unbiased sample).

Final weights. A 'pseudo-population' that has all confounding removed is created using the final weights by simply multiplying the individual weights.

Weighted outcome analysis

In the second stage the treatment-outcome relationship of interest is assessed using the re-weighted sample. Validity of marginal structural models is assessed with several sensitivity analyses. The distributions of treatment weights, censoring weights and final weights are usually assessed graphically. If extreme values are identified sensitivity analyses are conducted by comparing results of outcome analyses including and excluding outliers (see limitations section and alternative approaches). Often, however, marginal structural models require fitting several different variations of each of the weight-generating models to achieve optimal weight distributions.

DO BETA-BLOCKERS REDUCE MORTALITY POST-MYOCARDIAL INFARCTION?

Delaney *et al.* studied the effects of beta-blockers on the risk of death in a cohort of people discharged from hospital with a diagnosis of acute myocardial infarction [1]. Both beta-blocker use and blood pressure were identified in both the 90-day period before and the 90-day period after the myocardial infarction. Table 1 shows the results of different analytical approaches. Results of marginal structural models are similar to pooled summary effects from randomized controlled trials reported in a meta-analysis.

INTERPRETATION

Results of marginal structural models have similar interpretation as clinical trials (i.e. a marginal or population-level interpretation). Marginal structural models estimate what would happen if a person always received a certain treatment versus never, which is an idealized situation that does not reflect clinical practice, unless it is interpreted as an 'intention to continue treatment' similar to the 'intention to treat' interpretation of randomized controlled trials. Other methods that address time-varying confounding affected by previous treatment allow different types of inference based on a conditional, as opposed to marginal interpretation. For example, the sequential Cox approach (described below) estimates the effect of starting a treatment versus never, and ignoring previous treatment [16].

Table 1. Effects of use of beta-blockers (yes versus no) on mortality post-myocardial infarction

Modeling method	Rate ratio	95% confidence intervals
Crude model (no adjustment)	0.31	0.26–0.37
Regression model[a]	0.54	0.45–0.67
IPTW model[b]	0.72	0.61–0.84
Meta-analysis[c]	0.77	0.69–0.85

[a]Regression model including only baseline covariates.
[b]Inverse-probability-of-treatment weighting (IPTW) model taking into account time-varying blood pressure values.
[c]Meta-analysis of randomized controlled trials [1].

STRENGTHS AND LIMITATIONS

The main strength and added value of marginal structural models is that they do not suffer from collider stratification bias because weighting, as opposed to conditioning, is used to control for time-varying confounders affected by previous treatment. However, marginal structural models have limitations. First, marginal structural models (as all causal models) can only achieve balance on known factors, and the exchangeability assumption is not verifiable. Second, the number of balancing variables may be limited by sample size. Unusual (or very common) covariates histories may result in failure to achieve stability of the estimated weights. This is why sensitivity analyses are extremely important in marginal structural models. Consistency of results is usually assessed by trimming upper and lower percentiles (e.g. resetting <1st percentile = 1st percentile and >99th percentile = 99th percentile) or by including and excluding observations with extreme values. Third, standard IPTW-based marginal structural models need to include all covariates in the weight estimation. Interaction effect can be estimated for baseline modifiers but not for time-varying modifiers in standard marginal structural models (although history-adjusted marginal structural models have been formulated) [17].

ALTERNATIVE METHODS TO CONTROL FOR TIME-VARYING CONFOUNDING

Before marginal structural models were first formulated [8], other two approaches had been proposed to estimate the causal effect of a time-varying treatment in the presence of time-varying confounders that are affected by previous treatment (exposure): the G-computation formula [18] and G-estimation of structural nested models [19]. G-computation works by first modeling the relationships between observed data using standard approaches. Using these models, G-computation predicts, for each individual, the unobserved outcome associated with the alternative (counterfactual) treatment (exposure) [20]. G-computation estimates the treatment (exposure) effect as the difference between observed and counterfactual outcomes either as a single (marginal or average) effect estimate or covariate-specific (conditional) effect. G-estimation is a semi-parametric method to estimate the effects of treatment (exposure) in structural nested models. These models control for time-varying confounding affected by previous treatment (exposure) by modeling the outcome at each time as a function of the treatment and covariate history up to that time [17]. Although G-computation and G-estimation methods have advantages over marginal structural models (e.g. more flexible modeling of time-varying effect modification and robustness to situations that would threaten the positivity assumption of IPTW-based procedures) they remain under-used; partly for limited implementation in standard statistical packages [17]. A tutorial has been recently published on these methods [21].

Sequential Cox models are a simple and intuitive approach to estimate treatment effects in the presence of time-varying

confounding affected by previous treatment [16]. This method uses longitudinal data (including repeated measures) to mimic several randomized controlled trials. Each trial is constructed based on individuals starting treatment in a certain time interval. An overall effect estimate for all such trials is obtained with stratified Cox analysis on the joint data set of all trials, where each trial is one stratum. Although sequential Cox models are an alternative to marginal structural models, they may provide different estimates when applied to the same set of data (conditional as opposed to marginal), and require a different interpretation. They estimate the effect of starting (as opposed to receiving) treatment versus remaining untreated. Sequential Cox models allow adjustment for dependent censoring (using weights) and testing the potential modification effect of time (changing effect over time) or other variables, including time-varying covariates. One recent study applied this methodology to assess the role of high erythropoietin doses on mortality accounting for the time-varying confounding effect of anemia [22].

CONCLUSION

Marginal structural models are a method to control for the effect of confounding that changes over time and is affected by previous treatment. In these situations, the relationship between confounding variables and treatment (or exposure) is bi-directional and standard methods to estimate the relationship of interest between treatment and outcome are not appropriate. Several examples exist for the interested reader on how to implement these strategies using standard statistical packages, including R (https://www.r-project.org) and STATA (www.stata.com). The authors of the 'ipw' package in R use examples from the HIV literature to show how treatment weights are calculated and used to address the time-varying confounding of CD4 count on the effect of active anti-retroviral therapy [23]. Fewell *et al.* discuss the use of marginal structural models to control time-varying confounding of disability index on the effect of methotrexate in people with rheumatoid arthritis [24]. In all these and other longitudinal study designs, marginal structural models are a powerful method for confounding control.

SUPPLEMENTARY DATA

Supplementary data are available online at http://ndt.oxfordjournals.org.

ACKNOWLEDGEMENTS

The authors thank Dr Ronald Geskus Ph.D. (Academic Medical Centre, Amsterdam, The Netherlands), the current maintainer of the package IPW in R for providing additional information on the package, and Dr Jon Michael Gran Ph.D. (Department of Biostatistics, University of Oslo, Norway) for providing additional information on his work on sequential Cox models.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare and declare that the results presented in this paper have not been published previously in whole or part.

REFERENCES

1. Delaney JA, Daskalopoulou SS, Suissa S. Traditional versus marginal structural models to estimate the effectiveness of beta-blocker use on mortality after myocardial infarction. *Pharmacoepidemiol Drug Saf* 2009; 18: 1–6
2. Ravani P, Palmer SC, Oliver MJ *et al.* Associations between hemodialysis access type and clinical outcomes: a systematic review. *J Am Soc Nephrol* 2013; 24: 465–473
3. Grubbs V, Wasse H, Vittinghoff E *et al.* Health status as a potential mediator of the association between hemodialysis vascular access and mortality. *Nephrol Dial Transplant* 2014; 29: 892–898
4. Ravani P, Parfrey P, Gadag V *et al.* Clinical research of kidney diseases V: extended analytic models. *Nephrol Dial Transplant* 2008; 23: 1484–1492
5. Ravani P, Parfrey P, Murphy S *et al.* Clinical research of kidney diseases IV: standard regression models. *Nephrol Dial Transplant* 2008; 23: 475–482
6. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011; 46: 399–424
7. Tingley D, Yamamoto T, Hirose K *et al.* Mediation: R package for causal mediation analysis. *J Stat Softw* 2014; 59: 1–38
8. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; 11: 550–560
9. Wang O, Kilpatrick RD, Critchlow CW *et al.* Relationship between epoetin alfa dose and mortality: findings from a marginal structural model. *Clin J Am Soc Nephrol* 2010; 5: 182–188
10. Lertdumrongluk P, Streja E, Rhee CM *et al.* Dose of hemodialysis and survival: a marginal structural model analysis. *Am J Nephrol* 2014; 39: 383–391
11. Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* 2004; 58: 265–271
12. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008; 168: 656–664
13. Pearl J. On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology* 2010; 21: 872–875
14. Tilling K, Sterne JA, Szklo M. Estimating the effect of cardiovascular risk factors on all-cause mortality and incidence of coronary heart disease using G-estimation: the atherosclerosis risk in communities study. *Am J Epidemiol* 2002; 155: 710–718
15. Mortimer KM, Neugebauer R, van der Laan M *et al.* An application of model-fitting procedures for marginal structural models. *Am J Epidemiol* 2005; 162: 382–388
16. Gran JM, Roysland K, Wolbers M *et al.* A sequential Cox approach for estimating the causal effect of treatment in the presence of time-dependent confounding applied to data from the Swiss HIV Cohort Study. *Stat Med* 2010; 29: 2757–2768
17. Joffe MM. Structural nested models, g-estimation, and the healthy worker effect: the promise (mostly unrealized) and the pitfalls. *Epidemiology* 2012; 23: 220–222
18. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model* 1986; 7: 1393–1512
19. Robins JM, Blevins D, Ritter G *et al.* G-estimation of the effect of prophylaxis therapy for Pneumocystis carinii pneumonia on the survival of AIDS patients. *Epidemiology* 1992; 3: 319–336
20. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol* 2011; 173: 731–738

21. Daniel RM, Cousens SN, De Stavola BL *et al.* Methods for dealing with time-dependent confounding. *Stat Med* 2013; 32: 1584–1618
22. Suttorp MM, Hoekstra T, Mittelman M *et al.* Treatment with high dose of erythropoiesis-stimulating agents and mortality: analysis with a sequential Cox approach and a marginal structural model. *Pharmacoepidemiol Drug Saf* 2015; 24: 1068–1075
23. van der Wal WM, Geskus RB. An R package for inverse probability weighting. *J Stat Softw* 2011; 43: 1–23
24. Fewell Z, Hernan MA, Wolfe F *et al.* Controlling for time-dependent confounding using marginal structural models. *Stata J* 2004; 4: 402–420

Received for publication: 26.5.2016; Editorial decision: 26.8.2016