Causal modeling in epidemiological practice

van der Wal, W.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 9

General Discussion

In this thesis, the use of MSMs to estimate causal effects is studied in epidemiological practice. We have implemented the most widely used method to fit MSMs, IPW, in the R package ipw (chapter 2). We applied IPW to fit MSMs, estimating the causal effect of exposures on outcomes using observational longitudinal data in the context of nephrology (chapters 3 and 4), and HIV research (chapters 5 and 7). In chapter 6, it was originally planned to use a variant of IPW to adjust for informative censoring, IPCW. However, it was concluded that in this case, the use of a "standard" method, namely stratification on a baseline covariate, was adequate. In chapter 7, the use of an alternative method to fit MSMs, G-computation, was studied. We developed an algorithm to quantify the causal effect of a secondary illness on the progression of a chronic disease. In chapter 8, small sample bias of IPW due to practical violations of the assumption of positivity was studied. We developed an algorithm to estimate IPW weights iteratively, resulting in causal effect estimators that suffer substantially less from this small sample bias.

Below, we draw conclusions from this thesis, and identify challenges and opportunities for future research on causal modeling in epidemiology.

9.1 This thesis

Of the different methods that are available to fit MSMs (e.g. IPW, G-computation, DR-estimation and TMLE, IPW is the most easily implemented using standard software. To allow for the implementation of IPW in a specific statistical software package, it is necessary that 1) weights can be computed, e.g. using predicted probabilities from a logistic model, 2) data can be weighted by these weights in a subsequent analysis, fitting a regression model estimating the parameters of a MSM, and 3) inference can be performed in which the clustering that is present in the weighted dataset (see section 1.8.1) is taken into account, by using a robust standard error estimator such as a sandwich estimator (Huber, 1967),
by bootstrapping (Efron & Tibshirani, 1993), or when possible using the delta method (e.g. as was used by Hajat et al., 2010). It has to be studied whether it is necessary to take into account the possible uncertainty that is introduced in IPW estimators by estimating the weights, e.g by bootstrapping the entire procedure comprised of 1) and 2) above. In a simulation study that we originally designed to explore this question, IPW small sample bias due to practical non-positivity seemed to be more problematic than potentially poor coverage of confidence intervals (chapter 8). Our IIPW algorithm, that suffers substantially less from this small sample bias than standard IPW, seemed to have reasonably adequate coverage, without taking this uncertainty into account. Nevertheless, our R package ipw can be used both with sandwich variance estimators and with bootstrapping, in a wide range of practical situations, to fit MSMs on observational data.

As illustrated in chapter 3, the use of IPW allows for combining different types of corrections, for example to correct both for confounding of the effect of the exposure of interest at baseline, for confounding as a result of cross-over from one treatment arm to the other due to covariates that also affect the outcome, and for informative censoring. To do so, separate weightings can be combined by multiplying the weights. In chapter 3, we have taken into account possible uncertainty introduced by estimating these three types of weights, by bootstrapping the entire procedure. The MSM used was a Cox model regressing the hazard of death on the exposure (renal dialysis type). This model also included main effects of baseline covariates and follow-up time, including non-linear effects of the latter. Furthermore, the model included the interactions of the baseline covariates and follow-up time with the exposure. From this model, causal effects conditional on specific values of the baseline covariates and follow-up time could be produced. It was noticed that hazard ratios produced from such a complex model were not straightforward to interpret. Therefore, we recommend to compute cumulative survival from such models, as we have done. These are more easily interpreted, especially when depicted graphically. The resulting survival curves are comparable to the inverse probability weighted survival curves of Cole & Hernán (2004).

MSMs are most often used to quantify the effect of a medical treatment that is categorical, often dichotomous (e.g. the initiation of zidovudine treatment, Hernán et al. (2000)). However, in chapter 4 we have used IPW to fit a MSM that quantifies the causal effect of a continuous variable which is an indicator of the health status of a patient (residual renal function). We have described the technical details of using IPW to correct for confounding with residual renal function as continuous exposure, having a zero-inflated distribution. However, the exposure, residual renal function, is not a clearly defined intervention, which may lead to a violation of the assumption of consistency. A similar argument can be made for chapters 5 and 7 in which we have estimated the effect of the development of diseases (AIDS defining conditions and active tuberculosis, respectively). In

\footnote{See section 1.6, and see also Hernán & Robins (2011) for a thorough discussion of this subject.}
principle, it is possible to make the assumption of consistency for such an exposure, when the mechanism involved is fully known, and we have specified which aspects of this mechanism comprise the particular exposure for which we are trying to estimate a causal effect. In chapters 4, 5 and 7 we have assumed that the causal effects that were estimated respectively suggested that 1) preventing or delaying the full loss of residual renal function could lead to lower mortality in patients with ESRD and 2) preventing or delaying the development of AIDS defining conditions (including active tuberculosis) could lead to lower mortality in HIV positive individuals.

In chapter 5 we have estimated the causal effects of both AIDS defining conditions (ADCs) and highly active antiretroviral therapy (HAART) on mortality, conditional on calendar time, including interactions between ADCs, HAART and calendar time. The goal of this study was to explore the relationship between these three independent variables and the dependent variable time to death. A different solution to reach this goal is to only designate ADCs as exposure, and estimate its effect conditional on both HAART use and calendar time, using a history-adjusted MSM (HA-MSM). Similarly, a HA-MSM could be used to estimate the causal effect of HAART, conditional on observed ADCs and calendar time. HA-MSMs consist of multiple MSMs fitted on data within a chosen window of follow-up time for different time points. As we have found in our study, this approach has three important drawbacks: 1) By using a HA-MSM, it is assumed that the time-varying covariates remain constant during the remaining time in every window, which is in general unrealistic and can lead to bias due to misclassification; 2) Combining the separate MSMs for every time window is often necessary with a low number of events, and aids interpretation, but this can lead to a dataset so large that fitting it using a standard desktop computer is impractical because of memory limitations; 3) The results from a HA-MSM are conditional on the value of time-varying covariates at the start of every time window. E.g. using the HA-MSM approach we could estimate the causal effect of time-varying ADC status within every time window conditional on HAART use at the start of the window. Such an effect is difficult to interpret.

When adjusting for confounding or informative censoring, standard conditioning is at present the most widely used method in epidemiology. MSMs, e.g. fitted using IPW, present an alternative to standard conditioning, that does not suffer from the important drawbacks of conditioning (see section 1.7). However, it is important to distinguish between conditioning to correct for confounding (or informative censoring), and conditioning to estimate conditional effects. The latter is useful when it is of interest to estimate effects conditional on measured covariates, e.g. within subgroups based on gender. In the latter example, the conditional model should include both main effects of the exposure and gender, as well as the interaction between the two. Conditioning to estimate causal effects, conditional on the measured value of confounders, can be done in MSMs. Estimates from such a conditional MSM can be adjusted for other covariates than are conditioned on, e.g. by IPW. We have fitted such conditional causal models throughout this the-
sis. However, in chapter 6 we have used standard conditioning both to adjust for informative censoring and to examine subgroups in a longitudinal study of quality of life measurements. In this study it was not necessary to use IPCW, since the drawbacks of standard conditioning did not apply.

As noted, IPW may suffer from small sample bias due to practical violations of the positivity assumption. We have studied alternative methods to fit MSMs. In chapter 7 we describe an algorithm that uses G-computation to quantify the causal effect of a secondary illness on the progression of a chronic disease. G-computation is known to be robust against practical non-positivity (see section 1.8.2). The algorithm can be used in a specific practical setting, when latent baseline variables (e.g. a random intercept and slope) fully describe the “true” value of the time-varying confounders, separate from short term fluctuations and measurement error. However, in general with a time-varying exposure as well as time-varying confounders, G-computation remains very challenging to implement. The iterative algorithm to estimate IPW that we describe in chapter 8 suffers substantially less than standard IPW from small sample bias due to practical non-positivity, in point treatment situations. However, it is necessary to further study the performance of this algorithm in longitudinal settings. After generalizing IIPW to longitudinal settings, we plan to add it to our R package ipw (see chapter 2).

9.2 Future perspectives

Based on this thesis, and the available literature, important challenges and opportunities remain:

1. The small sample behaviour of causal effect estimators (both bias and efficiency) can be sub-optimal. In particular, IPW seems to have substantial small sample bias and is less efficient than standard conditional methods when using non-saturated models. When confounding is present, implying an effect of confounders on the exposure, practical non-positivity can always occur when samples are sufficiently small. This is unfortunate, since IPW is the most widely used method for fitting MSMs. Newer estimators, notably TMLE, seem to have better small sample behaviour, but are harder to implement. Our IIPW method seems promising but has to be generalized to longitudinal settings.

2. When using IPW, inference can be done using robust variance estimators, bootstrapping or the delta method. Coverage and length of CIs based on these methods should be studied, both theoretically and through simulations.

3. Currently available software to fit MSMs, including inference, is still limited, both in accessibility and in scope. Therefore, these methods are not
widely available to applied researchers. Software available for the analysis of longitudinal data is especially limited. Illustrations of program code are available, which have to be adapted to each new setting. There is a need for more generally usable software packages, to suit applied researchers that are not familiar with extensive statistical programming. Our \texttt{R} package \texttt{ipw} is easily accessible and very flexible. Therefore, inverse probability weighting is now in principle accessible to applied researchers. Still, such flexible software implementation is lacking for other methods.

4. With publications in epidemiological journals such as Epidemiology, or applied journals such as Kidney International, the use of MSMs seems to become more mainstream. However, MSMs are not fully disseminated yet. It remains a challenge to convince applied researchers of the necessity of the causal inference methodology. Preferably, master courses in causal inference should be developed.

5. The implications of using MSMs when estimating the effect of exposures other than medical treatments should be studied. E.g. for etiological reasons, it could be of interest to estimate the causal effect of a certain disease on a relevant outcome. The interpretation of such estimates should be studied in detail.

6. The use of MSMs in practice should be studied in different fields, each having its own specific problems. E.g. in cost benefit-analysis it is important to estimate marginal absolute risks, possibly adjusted for confounding, which are readily obtained from MSMs. In longitudinal public health databases, left-truncation of data is common. In such a situation, it is not straightforward to correct for time-varying confounders, when part of their history is missing.