Practice of Epidemiology

Clustering and Residual Confounding in the Application of Marginal Structural Models: Dialysis Modality, Vascular Access, and Mortality


* Correspondence to Dr. Jessica Kasza, Department of Epidemiology and Preventive Medicine, Monash University, The Alfred Center, 99 Commercial Road, Melbourne, VIC 3004, Australia (e-mail: jessica.kasza@monash.edu).

Initially submitted December 3, 2014; accepted for publication March 30, 2015.

In the application of marginal structural models to compare time-varying treatments, it is rare that the hierarchical structure of a data set is accounted for or that the impact of unmeasured confounding on estimates is assessed. These issues often arise when analyzing data sets drawn from clinical registries, where patients may be clustered within health-care providers, and the amount of data collected from each patient may be limited by design (e.g., to reduce costs or encourage provider participation). We compared the survival of patients undergoing treatment with various dialysis types, where some patients switched dialysis modality during the course of their treatment, by estimating a marginal structural model using data from the Australia and New Zealand Dialysis and Transplant Registry, 2003–2011. The number of variables recorded by the registry is limited, and patients are clustered within the dialysis centers responsible for their treatment, so we assessed the impact of accounting for unmeasured confounding or clustering on estimated treatment effects. Accounting for clustering had limited impact, and only unreasonable levels of unmeasured confounding would have changed conclusions about treatment comparisons. Our analysis serves as a case study in assessing the impact of unmeasured confounding and clustering in the application of marginal structural models.

clinical registries; clustering; dialysis; marginal structural model; unmeasured confounding

Abbreviations: ANZDATA, Australia and New Zealand Dialysis and Transplant; AVF, arteriovenous fistula; AVG, arteriovenous graft; MSM, marginal structural model; SIPTW, stabilized inverse probability of treatment weights.

Marginal structural models (MSMs) (1, 2) are useful tools for the estimation of treatment effects in situations where treatments may change over time and affect and be affected by time-varying confounders. However, clustering is only rarely accounted for in applications of MSMs and other causal methods (3, 4) and, although the assumption of no unmeasured confounders is often mentioned and may be discussed, the sensitivity of conclusions to this assumption is not often assessed.

Registries have the potential to lead to improvements in the quality of health-care delivery (5) by providing insights into treatment comparisons in situations where definitive clinical trials or appropriate-scale cohort studies are ethically indefensible or logistically difficult. However, the analysis of data drawn from clinical and disease registries may be particularly vulnerable to unmeasured confounding. To reduce costs, increase ease of data collection, and achieve greater compliance, registries generally aim to collect information on a limited number of variables and on a limited number of occasions per patient; consequently, they may be missing information on important confounders. Further, the clustered nature of patient-care episodes by health-care provider lends a natural hierarchy to data on a clinical registry.

The Australia and New Zealand Dialysis and Transplant (ANZDATA) Registry collects data from all patients receiving dialysis for end-stage kidney disease in Australia and New Zealand (6). All dialysis patients have a “parent” center, with providers and professionals who are clinically responsible for the treatments and outcomes of patients at that center. In consultation with their clinicians, patients may elect to undergo either peritoneal dialysis (which uses the patient’s peritoneum as dialysis membrane) or hemodialysis (which uses a
dialysis machine to circulate blood past an artificial dialysis membrane). For the vast majority of patients, both forms of treatment are available and technically possible. Patients may change dialysis modality at any time during the course of their treatment. Data are collected at dialysis initiation, at the time of treatment changes, and in end-of-year annual surveys. There is a lack of consensus on which form of dialysis is associated with better survival, and it appears that differences are likely time dependent (7, 8). Additionally, the location of hemodialysis (home or at a facility) and the method by which access to the blood circulation is obtained, the “vascular access” type, have been shown to be associated with mortality (9, 10). Randomized clinical trials to compare peritoneal dialysis and hemodialysis are unfeasible; previous attempts in the United States and the Netherlands failed because of slow recruitment and patients’ lack of willingness to be randomized (11, 12). Analyses of data from clinical registries such as the ANZDATA Registry are necessary to inform the debate and guide decision making.

Changes in dialysis modality over the course of a dialysis patient’s treatment may be both a cause and an effect of changes in patient characteristics, so we apply an MSM to disentangle the relationship between dialysis modality history and mortality in the ANZDATA Registry, as has been previously applied to dialysis modality comparisons (10, 13, 14). For the estimated MSM parameters to have valid causal interpretations, in addition to the correct specification of all models, the treatment version irrelevance, positivity, and conditional exchangeability (implying no unmeasured confounders) assumptions must be valid. Treatment version irrelevance requires that different versions of the same treatment give rise to the same potential outcomes (15, 16), and positivity requires that each patient be assignable to each treatment type and that all covariate/treatment combinations be represented in the data set (17, 18).

In this article, we describe how the impact of unmeasured confounding and clustering of patients on estimated MSM parameters may be assessed. Clustering will be seen to be interlinked with the positivity assumption. The application of our methods to the ANZDATA Registry serves as a case study in assessing the impact of these challenges of applying MSMs to compare treatments in a clinical registry, and we provide example Stata code online. Although our application is to clinical registry data, our methods are more broadly applicable to other types of observational data.

**THE ANZDATA REGISTRY AND MARGINAL STRUCTURAL MODELS**

**ANZDATA Registry**

At the start of dialysis, the ANZDATA Registry records initial dialysis modality, date of birth, sex, racial origin, comorbid conditions (diabetes mellitus, coronary artery, peripheral vascular, cerebrovascular, and chronic lung diseases), late referral to a nephrologist (<3 months before dialysis start), height, weight, smoking status, serum creatinine, primary renal disease, and center of treatment. Vascular access is recorded as a central venous catheter (“temporary vascular access”), native arteriovenous fistula (AVF), or arteriovenous graft (AVG). We consider data from all adult patients who commenced dialysis from October 1, 2003, until December 31, 2011.

Surveys took place on the prespecified dates March 31, 2004, December 31, 2004, and then yearly to December 31, 2011. Surveys recorded current dialysis treatment, comorbid conditions, and dry weight at last dialysis. Vascular access at first ever hemodialysis receipt was recorded and, for all patients on hemodialysis at any point during the survey period, vascular access in use at the last hemodialysis session was recorded.

A total of 22,266 adult patients initiated dialysis during the study period, and 20,191 patients remained after the exclusions detailed in Web Appendix 1, available at [http://aje.oxfordjournals.org/](http://aje.oxfordjournals.org/). During the course of dialysis treatment, 6,971 patients died, 2,966 received kidney transplants, and 267 recovered kidney function. Patients were considered to be informatively censored at the time of kidney transplant or regain of kidney function. Table 1 describes the patients at commencement of dialysis.

Changes in dialysis modality, location (home/facility), death, transplantation, and regain of kidney function were recorded on the dates that they occurred. Vascular access, comorbidities, and weight were recorded only at dialysis start and the yearly survey and were imputed at nonsurvey times. Because changes in comorbidities are often associated with clinical events that may necessitate a change in dialysis modality, comorbidities at nonsurvey times were imputed by carrying the last observation back to the time of the previous survey. Changes in comorbid status were then supposed to occur at survey times only. Weight at nonsurvey times was imputed by carrying the last observed values forward, to ensure that measurement preceded any changes in exposure. Because the dates of vascular access change were not recorded, to incorporate uncertainty about the timing of vascular access changes in the analysis, we imputed 50 sets of central venous catheter to AVF/AVG change times (method 1 in Web Appendix 2, Web Tables 1 and 2, and Web Figure 1).

We subclassified hemodialysis by home versus facility and further subclassified by vascular access. At day 90 of dialysis, there were few home hemodialysis patients with a central venous catheter ($n = 26$) or an AVG ($n = 21$), so patients were censored at the time of a switch to home hemodialysis central venous catheter, and AVF and AVG categories were combined (as “permanent vascular access”). The final set of treatments is $A = \text{home hemodialysis AVF/AVG, facility hemodialysis AVF/AVG, facility hemodialysis central venous catheter, peritoneal dialysis})$.

**Marginal structural models**

We used an MSM to estimate the causal effect of dialysis modality history on mortality. Death, censoring events, and changes in dialysis modality and time-varying covariates were assumed to occur at discrete time points $t = 1, \ldots , T$. We discretized patient dialysis treatment time into 90-day periods, and changes in comorbidities are often associated with clinical events that may necessitate a change in dialysis modality, comorbidities at nonsurvey times were imputed by carrying the last observation back to the time of the previous survey. Changes in comorbid status were then supposed to occur at survey times only. Weight at nonsurvey times was imputed by carrying the last observed values forward, to ensure that measurement preceded any changes in exposure. Because the dates of vascular access change were not recorded, to incorporate uncertainty about the timing of vascular access changes in the analysis, we imputed 50 sets of central venous catheter to AVF/AVG change times (method 1 in Web Appendix 2, Web Tables 1 and 2, and Web Figure 1).

We subclassified hemodialysis by home versus facility and further subclassified by vascular access. At day 90 of dialysis, there were few home hemodialysis patients with a central venous catheter ($n = 26$) or an AVG ($n = 21$), so patients were censored at the time of a switch to home hemodialysis central venous catheter, and AVF and AVG categories were combined (as “permanent vascular access”). The final set of treatments is $A = \text{home hemodialysis AVF/AVG, facility hemodialysis AVF/AVG, facility hemodialysis central venous catheter, peritoneal dialysis})$.

**Marginal structural models**

We used an MSM to estimate the causal effect of dialysis modality history on mortality. Death, censoring events, and changes in dialysis modality and time-varying covariates were assumed to occur at discrete time points $t = 1, \ldots , T$. We discretized patient dialysis treatment time into 90-day periods, with the treatment during a period defined as that in use for the majority of the period.

Using the usual MSM notation (19), $D_i(t) = 1$ if patient $i$ died at time $t$ and 0 otherwise, $A_i(t)$ is patient $i$’s treatment
Table 1. Characteristics of Patients by Dialysis Modality (Subclassified by Vascular Access) at Day 90, Australia and New Zealand Dialysis and Transplant Registry, 2003–2011

<table>
<thead>
<tr>
<th>Factor</th>
<th>Hemodialysis Home AVF/AVG (n = 357)</th>
<th>Hemodialysis Facility AVF/AVG (n = 7,414)</th>
<th>Peritoneal Dialysis (n = 6,665)</th>
<th>P Valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yearsc</td>
<td>50.1 (11.2)</td>
<td>62.3 (14.1)</td>
<td>61.7 (15.7)</td>
<td>60.2 (14.8)</td>
</tr>
<tr>
<td>Serum creatinine, µmol/Ld</td>
<td>644 (535–800)</td>
<td>624 (485–806)</td>
<td>666 (500–896)</td>
<td>608 (474–798)</td>
</tr>
<tr>
<td>Female</td>
<td>81 22.7</td>
<td>2,028 35.4</td>
<td>2,976 40.1</td>
<td>2,812 42.2</td>
</tr>
<tr>
<td>Smoker</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Current</td>
<td>40 11.2</td>
<td>695 12.1</td>
<td>1,064 14.4</td>
<td>919 13.8</td>
</tr>
<tr>
<td>Former</td>
<td>127 35.6</td>
<td>2,456 42.9</td>
<td>3,083 41.6</td>
<td>2,651 39.8</td>
</tr>
<tr>
<td>Never</td>
<td>190 53.2</td>
<td>2,578 45.0</td>
<td>3,267 44.1</td>
<td>3,095 46.4</td>
</tr>
<tr>
<td>Late referral</td>
<td>17 4.8</td>
<td>455 7.9</td>
<td>2,902 39.1</td>
<td>1,220 18.3</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NZ Maori/Pacific people</td>
<td>28 7.8</td>
<td>426 7.4</td>
<td>930 12.5</td>
<td>885 13.3</td>
</tr>
<tr>
<td>Aboriginal/TSI</td>
<td>2 0.6</td>
<td>492 8.6</td>
<td>694 9.4</td>
<td>334 5.0</td>
</tr>
<tr>
<td>Asian</td>
<td>20 5.6</td>
<td>333 5.8</td>
<td>484 6.5</td>
<td>744 11.2</td>
</tr>
<tr>
<td>White/other</td>
<td>307 86.0</td>
<td>4,478 78.2</td>
<td>5,306 71.6</td>
<td>4,702 70.5</td>
</tr>
<tr>
<td>Year of first dialysis</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2003</td>
<td>6 1.7</td>
<td>157 2.7</td>
<td>196 2.6</td>
<td>214 3.2</td>
</tr>
<tr>
<td>2004</td>
<td>37 10.4</td>
<td>652 11.4</td>
<td>733 9.9</td>
<td>761 11.4</td>
</tr>
<tr>
<td>2005</td>
<td>44 12.3</td>
<td>694 12.1</td>
<td>922 12.4</td>
<td>807 12.1</td>
</tr>
<tr>
<td>2006</td>
<td>54 15.1</td>
<td>690 12.0</td>
<td>929 12.5</td>
<td>965 14.5</td>
</tr>
<tr>
<td>2007</td>
<td>46 12.9</td>
<td>714 12.5</td>
<td>959 12.9</td>
<td>839 12.6</td>
</tr>
<tr>
<td>2008</td>
<td>38 10.6</td>
<td>721 12.6</td>
<td>1,040 14.0</td>
<td>914 13.7</td>
</tr>
<tr>
<td>2009</td>
<td>49 13.7</td>
<td>772 13.5</td>
<td>964 13.0</td>
<td>854 12.8</td>
</tr>
<tr>
<td>2010</td>
<td>47 13.2</td>
<td>713 12.4</td>
<td>976 13.2</td>
<td>751 11.3</td>
</tr>
<tr>
<td>2011</td>
<td>36 10.1</td>
<td>616 10.8</td>
<td>695 9.4</td>
<td>560 8.4</td>
</tr>
<tr>
<td>Primary renal disease</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>64 17.9</td>
<td>2,037 35.6</td>
<td>2,852 38.5</td>
<td>2,385 35.8</td>
</tr>
<tr>
<td>Glomerulonephritis</td>
<td>142 39.8</td>
<td>1,256 21.9</td>
<td>1,516 20.4</td>
<td>1,654 24.8</td>
</tr>
<tr>
<td>Hypertension</td>
<td>23 6.4</td>
<td>782 13.6</td>
<td>967 13.0</td>
<td>900 13.5</td>
</tr>
<tr>
<td>Other</td>
<td>128 35.9</td>
<td>1,654 28.9</td>
<td>2,079 28.0</td>
<td>1,726 25.9</td>
</tr>
<tr>
<td>Body mass indexe</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>15–19.9</td>
<td>11 3.1</td>
<td>312 5.4</td>
<td>624 8.4</td>
<td>484 7.3</td>
</tr>
<tr>
<td>20–24.9</td>
<td>79 22.1</td>
<td>1,505 26.3</td>
<td>2,173 29.3</td>
<td>2,095 31.4</td>
</tr>
<tr>
<td>25–29.9</td>
<td>137 38.4</td>
<td>1,847 32.2</td>
<td>2,243 30.3</td>
<td>2,349 35.2</td>
</tr>
<tr>
<td>30–49.9</td>
<td>130 36.4</td>
<td>2,065 36.0</td>
<td>2,374 32.0</td>
<td>1,737 26.1</td>
</tr>
<tr>
<td>Comorbid conditions</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>64 17.9</td>
<td>2,392 41.8</td>
<td>3,461 46.7</td>
<td>2,469 37.0</td>
</tr>
<tr>
<td>Lung disease</td>
<td>25 7.0</td>
<td>919 16.0</td>
<td>1,451 19.6</td>
<td>1,040 15.6</td>
</tr>
<tr>
<td>Diabetes type 1</td>
<td>10 2.8</td>
<td>150 2.6</td>
<td>223 3.0</td>
<td>292 4.4</td>
</tr>
<tr>
<td>Diabetes type 2</td>
<td>75 21.0</td>
<td>2,514 43.9</td>
<td>3,508 47.3</td>
<td>2,673 40.1</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>36 10.1</td>
<td>1,441 25.2</td>
<td>2,189 29.5</td>
<td>1,562 23.4</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>13 3.6</td>
<td>835 14.6</td>
<td>1,293 17.4</td>
<td>961 14.4</td>
</tr>
</tbody>
</table>

Abbreviations: AVF, arteriovenous fistula; AVG, arteriovenous graft; CVC, central venous catheter; NZ, New Zealand; TSI, Torres Strait Islander.

a Twenty-six patients with home hemodialysis CVC at day 90 were excluded.

b P value from analysis of variance or χ² test comparing patient characteristics between dialysis modality/vascular access categories.

c Age is presented as mean (standard deviation).

d Serum creatinine is presented as median (interquartile range).

e Weight (kg)/height (m)².
at time $t$, $A_i(t) \in A_i$, and $V_i$ is patient $i$’s set of baseline covariates consisting of all covariates listed in Table 1, including the baseline levels of the time-varying comorbidities and body mass index. The set of time-varying covariates for patient $i$ at time $t$, consisting of comorbidity indicators (excluding diabetes status, as only 120 patients developed diabetes) and body mass index, is denoted by $L_i(t)$.

To estimate the causal effect of dialysis modality history on mortality, we fit a pooled logistic regression model to the data:

$$\logit[P(D_i(t)|D_i(t-1) = 0, A_i(t), V_i)] = \beta_0(t) + \beta_1(t)A_i(t) + \beta_2V_i,$$

where each patient’s observation at each time period $t$ was weighted by his/her stabilized inverse probability of treatment and censoring weight. We assumed that mortality depended on dialysis history only through current dialysis modality. If time intervals are short with a small probability of the occurrence of an event in each interval, the odds ratios of this model are equivalent to the hazard ratios of a marginal structural Cox proportional hazards model (20). Hence, we refer to the odds ratios in the model as hazard ratios. Natural cubic spline terms for period, as well as interaction terms between treatment and these spline terms, were included in the model to allow for nonproportional hazards over time, as indicated by the dependence of $\beta_0$ and $\beta_1$ on $t$ in equation 1.

The stabilized inverse probability of treatment and censoring weights are the products of the stabilized inverse probability of treatment weights (SIPTWs) and the stabilized inverse probability of censoring weights. Stabilized weights were used to reduce the variability of the weights and the standard errors of the estimated hazard ratios (21). The SIPTW for patient $i$ at period $t$ is given by the following:

$$\text{SIPTW}_i(t) = \prod_{s=1}^t \frac{P(A_i(s) = a_i(s)|V_i, A_i(s-1))}{P(A_i(s) = a_i(s)|V_i, A_i(s-1), L_i(s-1))}.$$ 

Multinomial regression models for current treatment were used to estimate probabilities. Natural cubic splines for period were included to allow for changes in the intercept over time.

Stabilized inverse probability of censoring weights were estimated similarly to SIPTWs, with dependence on treatment and baseline and time-varying covariates, using logistic regression models for censoring events. Separate MSMs were fitted for each of the 50 sets of imputed vascular access change times, and estimates were used by using Rubin’s rules as described by Carpenter and Kenward (22). All analyses used Stata, release 12, software (23). Computational details for similar models are available as given elsewhere (24).

Figure 1 displays the estimated time-dependent hazard ratios for mortality for each of facility hemodialysis with a central venous catheter, peritoneal dialysis, and home hemodialysis with an AVF/AVG, relative to facility hemodialysis with an AVF/AVG.

**METHODS**

**Clustering by dialysis center and the positivity assumption**

The 20,191 patients included in the final analysis were clustered within 85 dialysis centers. Treatment modality and patterns of vascular access use vary across centers as do survival outcomes. Patient characteristics explain some of this variation, but much remains unexplained, with many potentially important but unmeasured characteristics likely to vary between centers (which may be related to the selection and implementation of each treatment type), so it is necessary to account for clustering in the estimation of these models. Center effects could be included as fixed or random in all models, but since the aim of MSMs is to estimate the population average treatment effect, including fixed center effects in the survival model is preferred (3). However, there is no such clear choice for the treatment and censoring models. We include fixed effects for center in the treatment and censoring models to reduce computation time when fitting models for 50 imputations of vascular access change times.

To satisfy the positivity assumption, it is necessary that all treatments be available in all clusters. Web Figure 2 displays the proportions of 90-day periods attributed to each of the treatments within each center, with centers labeled with and ordered by the total number of 90-day periods at that center, for 1 imputation of vascular access change times. There were some small centers in which only facility hemodialysis AVF/AVG was represented and some larger centers in which some treatments were never or rarely used, although in general, as center size increases, the treatment mix becomes more varied.

We denote the set of treatments available at center $C_i$ by $\mathcal{A}_i$. If treatment were binary, to satisfy the positivity assumption, data from all centers that used only 1 treatment must be excluded from the analysis. In our situation with multicategory conditions.
treatment, data from centers with \( A_j \subset A \) and \(|A_j| \geq 2\) can be included in analyses by weakening the positivity assumption. For patient \( i \) in \( C_j \), the weakened positivity assumption requires

\[
\frac{P(A_i(t) = a|y_i(t-1), V_{ij}, C_j)}{P(A_i(t) = a|y_i(t-1), D_y(t-1) V_{ij}, C_j)} < \infty, \quad \forall a_i \in A_j.
\]

The weakened positivity assumption has implications for the interpretation of estimates: Hazard ratios for particular treatments are defined only for centers in which the treatment option is available. A similar restriction of the treatment set (for the entire sample rather than for specific treatments) has been described previously (25).

For prevention of structural violations of the positivity assumption, instead of including all treatments that are possible within a center, we restricted the set of available treatments within a center \( C_j \) to \( A_j = \{a \in A| P(A = a|C_j) > \alpha \} \), with \( \alpha > 0 \). When \( \alpha > 0 \), patients are censored at the commencement of an unlikely treatment option. We consider \( \alpha = 0 \) or 0.05 and thus exclude impossible/unlikely center-treatment pairings.

When there are centers \( j \) with \( A_j \subset A \), the use of the multinomial logistic regression model to estimate SIPTWs is inappropriate, because such a model assigns non-0 probability to all treatments. The alternative-specific conditional logit model (also known as McFadden’s choice model (26)) with the available treatments as the alternatives, without any alternative-specific regressors, can be used to ensure that only \( a \in A_j \) are assigned non-0 probability.

**Assessing the impact of unmeasured confounding using a confounding function**

Because of the limited number of recorded covariates in the ANZDATA Registry and in registries more generally, negative control outcomes and instrumental variable approaches (27, 28) are not feasible options for assessing the impact of unmeasured confounding. If it is plausible to assume the existence of 1 or more unmeasured binary confounders, methods such as those described by Groenwold et al. (29) and Vanderweele et al. (30) could be applied to assess the impact of these particular hypothesized confounders on conclusions. However, such methods do not address the potential impact of unmeasured confounding in its entirety and require strong assumptions about the nature of the unmeasured confounders, which may be as untenable as the original assumption of no unmeasured confounding. The application of such methods when hypothesized confounders are time varying appears to be complex.

We modified an existing approach to assess the impact of unmeasured confounding, which makes no assumptions about individual unmeasured confounders (31, 32). This approach instead requires using contextual knowledge to quantify the effect of unmeasured confounding using a confounding function specified in terms of counterfactual outcomes.

The causal hazard ratio for treatment level \( a \in A \) (relative to baseline treatment level \( 0 \in A \)) is given by \( h_a(t) = \frac{P(D_a(t) = 1|D(t-1) = 0, V = v)P(D_0(t) = 1|D(t-1) = 0, V = v)}{P(D(t) = 1|D(t-1) = 0, A(t) = 0, V = v)} \) and is estimated by

\[
\hat{h}_a(t) = \frac{P(D(t) = 1|D(t-1) = 0, A(t) = a, V = v)}{P(D(t) = 1|D(t-1) = 0, A(t) = 0, V = v)}.
\]

The confounding function

\[
c(a, v, t) = \frac{P(D_a(t) = 1|D(t-1) = 0, V = v)}{\sum_{a^* \in A \setminus \{a\}} P(D_a(t) = 1|D(t-1) = 0, A(t) = a^*, V = v)P(A(t) = a^*|V = v)}
\]

quantifies the impact of unmeasured confounding on \( \hat{h}_a(t) \) at period \( t \), given baseline variables \( v \), where \( P(D_a(t) = 1|D(t-1) = 0, A(t) = a^*, V = v) \) is a counterfactual probability: the probability of death among patients with baseline characteristics \( V = v \), on treatment \( a^* \), had these patients actually received treatment \( a \). For example, for facility hemodialysis with a central venous catheter, the confounding function is the probability of death among patients actually receiving facility hemodialysis with a central venous catheter, given baseline characteristics, divided by a weighted average over all patients on other treatments of the (estimated) probability of death among those patients had they, contrary to fact, been receiving facility hemodialysis central venous catheter at \( t \), given baseline characteristics. Each term in the sum is weighted by the (estimated) probability of receiving the actual treatment (again, given baseline characteristics).

Assuming levels of confounding given by \( c(a, v, t) \) for \( a \in A \), the bias in the estimator \( \hat{h}_a(t) \) due to unmeasured confounding is

\[
\hat{h}_a(t) \left[ 1 - \frac{P(A(t) = a|V = v) + (1/c(a, v, t)) \sum_{a^* \in A \setminus \{a\}} P(A(t) = a^*|V = v)}{P(A(t) = 0|V = v) + (1/c(0, v, t)) \sum_{a^* \in A \setminus \{0\}} P(A(t) = a^*|V = v)} \right].
\]

Informally, \( c(a, v, t) \geq 1 \) can be interpreted as a discrete-time hazard ratio, which aids in the selection of plausible values. Values of \( c(a, v, t) > 1 \) imply that some of the observed risk of death is due to some unmeasured ill health. There is a greater risk of death for

those patients assigned to \( a \) than on average among those patients not assigned to \( a \) had they been assigned to \( a \); \( c(a, v, t) = 1 \) implies no difference in the risk of death of patients on \( a \) and those not on \( a \).

We recommend assessing the sensitivity of conclusions to various choices of confounding functions \((31)\). Depending on the application, it may be reasonable to assume time-invariance of the impact of unmeasured confounding, although even then, choosing appropriate values for confounding functions is not straightforward. For each treatment type, the impact of unmeasured confounding for a range of \( c(a, v, t) \) values can be assessed, a prior distribution could be placed on \( c(a, v, t) \), or, to help determine plausible magnitudes of unmeasured confounding, the impact of including additional variables in the MSM can be assessed.

An alternative approach is to calculate the \( c(a, v, t) \) value that is required to fully explain the estimated treatment effect if the true treatment effect were null (hazard ratio = 1). This \( c(a, v, t) \) value is defined by the following:

\[
c^I(a, v, t) = \frac{\hat{h}_a(t)(1 - P(A(t) = a | V = v))}{P(A(t) = 0 | V = v) + (1/c(0, v, t))(1 - P(A(t) = 0 | V = v)) - \hat{h}_a(t)P(A(t) = a | V = v)}.
\]

We provide example Stata code in Web Appendix 3 to account for clustering by provider and perform sensitivity analyses for unmeasured confounding online.

**RESULTS**

**Clustering by dialysis center**

To assess the impact of accounting for clustering by dialysis center on estimated hazard ratios, we fit models for the variety of clustering and treatment-set choices described in Table 2. Eleven centers with <150 treatment periods each were excluded (545 (0.3%) periods excluded) to avoid collinearities in the stabilized inverse probability of treatment and censoring weights and survival models.

Web Figure 3 displays estimated hazard ratios and confidence intervals at months 3, 6, and 12 of dialysis for each of the choices described in Table 2. For comparison, estimated hazard ratios from the MSM not accounting for clustering but fit to the data excluding the 11 centers are also displayed. For the most part, the estimated hazard ratios for the different treatment options changed very little when fixed center effects were included in the models. Unsurprisingly, given that they were based on the smallest number of centers and periods, hazard ratios from models C and D (refer to Table 2 for definitions) had the widest confidence intervals. For facility hemodialysis with a central venous catheter, the hazard ratios from models C and D were closer to 1 than for any other choices. Unlike for peritoneal dialysis and facility hemodialysis with a central venous catheter, the hazard ratios for home hemodialysis with an AVF/AVG for models G and H were somewhat different from those for models E and F. This is due to there being a number of centers with small but non-0 proportions of home hemodialysis AVF/AVG periods. The hazard ratios for models E–H, which used alternative-specific logit models to estimate SIPTWs, were generally close to the estimates obtained from the original model and models A (for models without fixed effects) and B (for models with fixed effects).

**Unmeasured confounding**

To help determine the potential magnitude of confounding functions for the ANZDATA Registry, we investigated the impact of including laboratory measurements in an MSM, as described in Web Appendix 4. The inclusion of laboratory measurements changed the hazard ratios minimally (Web Figure 4). Although this provides some reassurance that the impact of unmeasured confounding may be minor, that there

<table>
<thead>
<tr>
<th>Choice</th>
<th>Fixed Effects for Center*</th>
<th>Description of Data Set Included in Analysis</th>
<th>Formula for Data Set Included in Analysis</th>
<th>No. of Centersb</th>
<th>No. of Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>Centers with all treatments possible</td>
<td>( {C_r} P(A = a</td>
<td>C_r &gt; 0) \forall a \in \mathcal{A} )</td>
<td>68</td>
</tr>
<tr>
<td>B</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>No</td>
<td>Centers with all treatments probable</td>
<td>( {C_r} P(A = a</td>
<td>C_r &gt; 0.05) \forall a \in \mathcal{A} )</td>
<td>34</td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>No</td>
<td>For each center, only those treatments that are possible</td>
<td>( {a \in \mathcal{A}</td>
<td>P(A = a</td>
<td>C_r &gt; 0) } )</td>
</tr>
<tr>
<td>F</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>No</td>
<td>For each center, only those treatments that are probable</td>
<td>( {a \in \mathcal{A}</td>
<td>P(A = a</td>
<td>C_r &gt; 0.05) } )</td>
</tr>
<tr>
<td>H</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AVF, arteriovenous fistula; AVG, arteriovenous graft; CVC, central venous catheter; HD, hemodialysis; PD, peritoneal dialysis.

* If included, fixed effects for center are included in the treatment, censoring, and survival models.

b Numbers of centers (center / denoted by \( C _ r \)) and treatments included in each model are given for 1 simulation of vascular access change times. Centers with only 1 possible/probable treatment option are excluded from all analyses.

c The set of treatments is given by \( \mathcal{A} = \{ \text{home HD AVF/AVG}, \text{facility HD AVF/AVG}, \text{facility HD CVC, PD} \} \).
may be other entirely unmeasured quantities (e.g., mobility, social connectedness, or cognitive function) that have a stronger impact cannot be ruled out.

Figure 2 displays the results of our unmeasured confounding sensitivity analyses. We assumed for facility hemodialysis AVF/AVG patients (the most common and the reference exposure) that \( c(a, v, t) = 1 \). For each of facility hemodialysis with a central venous catheter, peritoneal dialysis, and home hemodialysis with an AVF/AVG, the left column of Figure 2 displays the estimated hazard ratio and the hazard ratio adjusted for varying amounts of time-invariant confounding. We supposed that facility hemodialysis central venous catheter patients were less healthy (given all measured covariates) than other patients, with \( c(\text{facility hemodialysis central venous catheter, } v, t) > 1 \). This implies that some of the observed worse survival of facility hemodialysis central venous catheter patients is due to some worse level of health, so corrected hazard ratios are shifted down toward 1. The assumption of time-invariant unmeasured confounding may be questionable, given the increase in the estimated hazard ratio for the home hemodialysis AVF/AVG group and the decrease in the facility hemodialysis central venous catheter hazard ratio over the first 24 months of dialysis followed by relatively stable hazard ratios (Figure 1), which may be indicative of the increasing similarity of dialysis patients over time.

The middle column of Figure 2 displays 2 choices of time-varying confounding functions and \( c^2(a, v, t) \), with hazard ratios corrected for these choices in the right column. The confounding functions for facility hemodialysis AVF/AVG patients imply that patients on this treatment have a greater risk of death than patients on other dialysis modalities initially, but that this risk declines over time. The impact of unmeasured confounding associated with peritoneal dialysis patients...
is supposed to be initially associated with a doubling/halving of the risk of death for peritoneal dialysis patients, with the impact diminishing over time. Home hemodialysis AVF/AVG patients are supposed to have a lower initial risk of death (given all measured confounders), increasing over time.

\( c^t(\text{peritoneal dialysis}, \nu, t) \) indicates that, in order for the difference between the estimated hazard ratio for peritoneal dialysis and 1 to be entirely due to unmeasured confounding, peritoneal dialysis patients must have more than twice the risk of death than patients assigned to other modalities, given measured confounders.

DISCUSSION

Clustering of patients within health-care providers and the impact of unmeasured confounding on estimates are not commonly addressed in the application of MSMs. Such issues are likely to arise when MSMs are applied to registry data, and given the growing importance of such applications in informing debate in situations in which clinical trials may be difficult or impossible, it is necessary to address these issues. We have demonstrated how these issues may be dealt with by using the ANZDATA Registry. Although there was little difference between the estimates for the various clustering approaches for our application, probably because the treatment mix was reasonably homogenous across centers, we recommend alternative-specific conditional logit models when available treatments differ across centers. In our analysis of the impact of unmeasured confounding, the magnitude of the impact as indicated by \( c^t(\text{peritoneal dialysis}, \nu, t) \) seems unlikely. Hence, it is plausible that the hazard ratio for peritoneal dialysis patients must be greater than 1 (the question remains, how much greater?). Similarly, the extreme values of \( c^t(\text{facility hemodialysis central venous catheter}, \nu, t) \) and \( c^t(\text{home hemodialysis AVF/AVG}, \nu, t) \) for \( t < 12 \) suggest that there must be some inflation of the risk of death for patients on facility hemodialysis with a central venous catheter and some reduction in the risk of death for home hemodialysis with an AVF/AVG, in the first year of dialysis.

Although the assumption of no unmeasured confounders is questionable in most applications, the sensitivity of conclusions to this assumption is rarely assessed. This may be due to the difficulty in applying established methods to situations where the treatment of interest is time varying or due to an inability to make assumptions about particular potential unmeasured confounders. The demonstrated approach does not require assumptions about particular confounders, and it can accommodate hypothesized time-varying impacts of unmeasured confounding.

The choice regarding dialysis modality is a significant treatment milestone for clinicians and their patients with end-stage kidney disease. Such decision making is based on a number of factors, including expected survival, independence, and quality of life. In the absence of definitive clinical trials, it is possible only to inform these decisions with observations. Our analyses suggest that peritoneal dialysis is associated overall with a higher rate of mortality than hemodialysis with an AVF/AVG in Australia and New Zealand. This result is robust to unmeasured confounding within a plausible range. The result at any individual center may vary from the population-averaged result, but in any case the relevance to individual patients will vary depending on the setting. For example, recent data suggest that patients will trade life expectancy for flexibility in treatments in their decisions regarding dialysis modality (33). Our analysis provides support for the use of MSMs for observational dialysis data under the assumption of no unmeasured confounding and provides robust information upon which clinicians and patients can base treatment decisions.

In our analysis, we assumed that patients did not change dialysis center during treatment. However, patients may change centers, and changes are recorded by the ANZDATA Registry. When all dialysis patients were considered, 8.8% had changes in dialysis center after day 90 of dialysis. Patients may change center to get access to a desired treatment, and changes in treatment coincident with changes in center, which accounted for 31% (788 of 2,519) of all center changes, may go toward explaining the existence of treatment-center pairings with small probabilities of occurring. However, the ANZDATA Registry does not collect reasons for center changes. Further work is required to determine the best way to account for center switches. We have dealt with the issues of unmeasured confounding and clustering separately here. If the effect of unmeasured confounding is thought to vary by center, considering the 2 issues separately would not be appropriate.

We have assumed that mortality is caused primarily by current dialysis modality, rather than by a combination of prior and current modalities. This assumption may not be valid. However, summarizing past dialysis history is difficult, and further work is required to determine how dialysis history may be more appropriately accounted for in MSMs.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria, Australia (Jessica Kasza, Kevan R. Polkinghorne, Rory Wolfe); ANZDATA Registry, Royal Adelaide Hospital, Adelaide, South Australia, Australia (Kevan R. Polkinghorne, Stephen P. McDonald); Departments of Nephrology and Medicine, Monash Medical Centre, Monash University, Melbourne, Victoria, Australia (Kevan R. Polkinghorne); Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand (Mark R. Marshall); Baxter Healthcare, Ltd., Panmure, Auckland, New Zealand (Mark R. Marshall); Department of Renal Medicine, Counties Manukau Health, Auckland, New Zealand (Mark R. Marshall); and Department of Medicine, University of Adelaide, Adelaide, South Australia, Australia (Stephen P. McDonald).

J.K. is supported by Australian National Health and Medical Research Council Centre of Excellence grant 1035261, awarded to the Victorian Centre for Biostatistics (ViCBiostat). The ANZDATA Registry is funded by the Australian Organ and Tissue Donation and Transplantation Authority, the New Zealand Ministry of Health, and Kidney Health Australia.

R.W. has received personal fees from Baxter Healthcare outside the submitted work, and M.R.M. is the Director of Medical Affairs for Baxter-Gambro Renal, Asia-Pacific.
There are no other relationships, activities, or support from any organization for any of the authors that could appear to have influenced the submitted work.

REFERENCES