Marginal structural models for clustered data: the positivity and no unmeasured confounding assumptions

Jessica Kasza, Kevan Polkinghorne, Mark R Marshall,
Stephen McDonald, Rory Wolfe

Department of Epidemiology and Preventive Medicine,
Monash University
ViCBiostat: Victorian Centre for Biostatistics

Outline

1. Motivating example: dialysis and the ANZDATA Registry
2. Clustering by dialysis centre
3. Unmeasured confounding and clustering

Outline

1. Motivating example: dialysis and the ANZDATA Registry
2. Clustering by dialysis centre
3. Unmeasured confounding and clustering

Dialysis, death, and ANZDATA

Question: which dialysis treatment modality is associated with longest survival times?

- Haemodialysis (HD)
  - **Home HD**: performed by the patient at home;
  - **Facility HD**: performed in a hospital/dialysis centre.
  - Vascular access (VA) types:
    - Arterio-venous fistula or graft: **AVF**/**AVG**
    - Central venous catheter: **CVC**
- **Peritoneal dialysis** (PD)

ANZDATA: Australia and New Zealand Dialysis and Transplant Registry

- Collects data from all dialysis patients in Australia and NZ.
- Changes between treatment modalities recorded as they occur.
- Data (including comorbidities, vascular access) collected at dialysis start and at yearly surveys.

The ANZDATA dataset used for analysis

All patients commencing dialysis between October 1 2003 and December 31 2011, undergoing at least 90 days of dialysis.

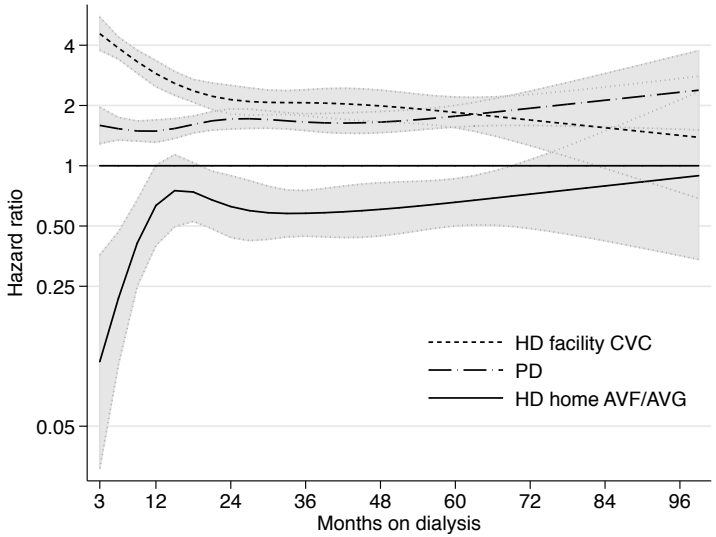
20,191 patients:

- 210,741 90-day periods of follow-up
- 6,971 deaths
- 2,966 kidney transplants
- 267 recovered kidney function

Over their treatment course, **30% of all patients had changes** in dialysis modality/VA

- Modality/VA choice thought to be affected by, and affect, comorbidities (e.g. coronary artery disease).
- We use marginal structural models (MSMs) to estimate the effect of modality/VA on mortality.

Estimated HRs, relative to facility HD AVF/AVG



MSM assumptions

- Consistency (treatment version irrelevance);
- Positivity: all patients have a positive probability of receiving all treatments;
- Conditional exchangeability (implying no unmeasured confounding).

MSM assumptions

- Consistency (treatment version irrelevance);
- Positivity: all patients have a positive probability of receiving all treatments;
- Conditional exchangeability (implying no unmeasured confounding).

Problems:

1. Patients are **clustered** within dialysis centres.
2. ANZDATA is a registry, so set of **measured confounders is limited**. Furthermore, the impact of unmeasured confounding may differ across clusters.

How do these problems impact upon validity of causal inference assumptions?

Outline

1. Motivating example: dialysis and the ANZDATA Registry

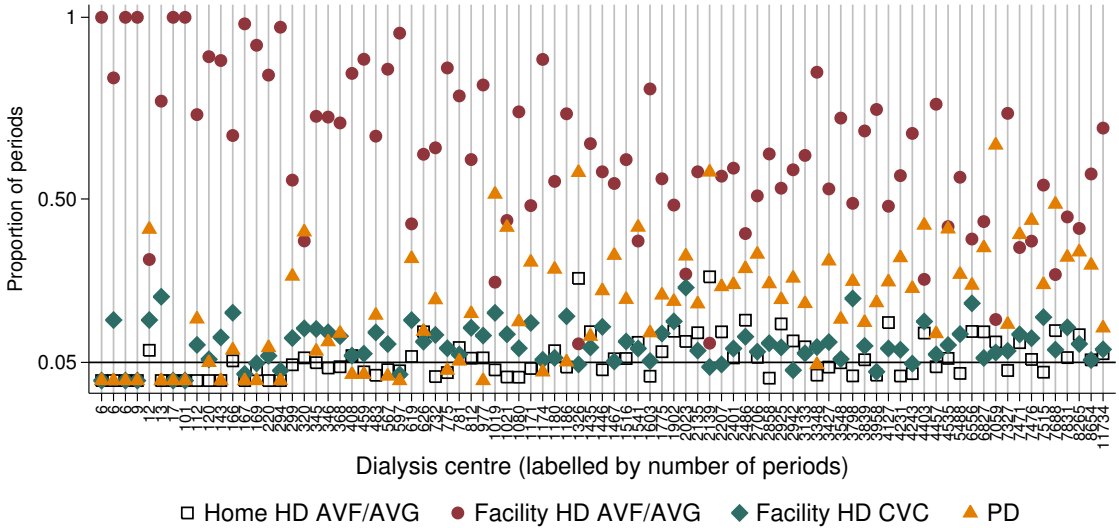
2. Clustering by dialysis centre

3. Unmeasured confounding and clustering

Clustering by dialysis centre

- All patients have a dialysis centre which is responsible for administering their treatment.
  - 85 dialysis centres are represented in our dataset.
  - There are differences in practice and survival across centres.
- An extreme difference: **not all dialysis types are available/represented in all centres** (or occur rarely within a centre).
  - In violation of the positivity assumption...

Clustering of treatments within the 85 centres



Clustering and the positivity assumption

Positivity: patient i at time t, on treatment $A_i(t) = a \in \mathcal{A}$, with baseline covariates V and time-varying covariates L

$$\frac{P(A_i(t) = a | A_i(t-1), V_i)}{P(A_i(t) = a | A_i(t-1), L_i(t-1), V_i)} < \infty \quad \forall a \in \mathcal{A}$$

Clustering and the positivity assumption

Positivity: patient i at time t, on treatment $A_i(t) = a \in \mathcal{A}$, with baseline covariates V and time-varying covariates L

$$\frac{P(A_i(t) = a | A_i(t-1), V_i)}{P(A_i(t) = a | A_i(t-1), L_i(t-1), V_i)} < \infty \quad \forall a \in \mathcal{A}$$

Cluster positivity: treatments available at centre $\mathcal{C}_j$ denoted by $a_k \in \mathcal{A}_j$

$$\frac{P(A_{ij}(t) = a_k | A_{ij}(t-1), V_{ij}, \mathcal{C}_j)}{P(A_{ij}(t) = a_k | A_{ij}(t-1), L_{ij}(t-1), V_{ij}, \mathcal{C}_j)} < \infty \quad \forall a_k \in \mathcal{A}_j$$

Clustering and the positivity assumption

Positivity: patient i at time t, on treatment $A_i(t) = a \in \mathcal{A}$, with baseline covariates V and time-varying covariates L

$$\frac{P(A_i(t) = a | A_i(t-1), V_i)}{P(A_i(t) = a | A_i(t-1), L_i(t-1), V_i)} < \infty \quad \forall a \in \mathcal{A}$$

Cluster positivity: treatments available at centre $\mathcal{C}_j$ denoted by $a_k \in \mathcal{A}_j$

$$\frac{P(A_{ij}(t) = a_k | A_{ij}(t-1), V_{ij}, \mathcal{C}_j)}{P(A_{ij}(t) = a_k | A_{ij}(t-1), L_{ij}(t-1), V_{ij}, \mathcal{C}_j)} < \infty \quad \forall a_k \in \mathcal{A}_j$$

If only cluster positivity holds, HRs for particular treatments only defined for centres in which treatment option is available.

Satisfaction of (cluster) positivity

Positivity: Restrict the set of centres

- Include only those centres in which all treatments are possible
- To prevent structural violations, include only centres in which all treatments occur at least 5% of the time (all *probable* treatments).

Satisfaction of (cluster) positivity

Positivity: Restrict the set of centres

- Include only those centres in which all treatments are possible
- To prevent structural violations, include only centres in which all treatments occur at least 5% of the time (all *probable* treatments).

Cluster positivity: Restrict the set of treatments within each centre

- Limit the sets of treatments at each centre to those possible/probable.
- Prevent unavailable treatments being assigned non-zero probabilities.

Satisfaction of (cluster) positivity

Positivity: Restrict the set of centres

- Include only those centres in which all treatments are possible
- To prevent structural violations, include only centres in which all treatments occur at least 5% of the time (all *probable* treatments).

Cluster positivity: Restrict the set of treatments within each centre

- Limit the sets of treatments at each centre to those possible/probable.
- Prevent unavailable treatments being assigned non-zero probabilities.

Both approaches: to account for unexplained variation between centres, include fixed effects for centres in treatment, censoring and survival models.

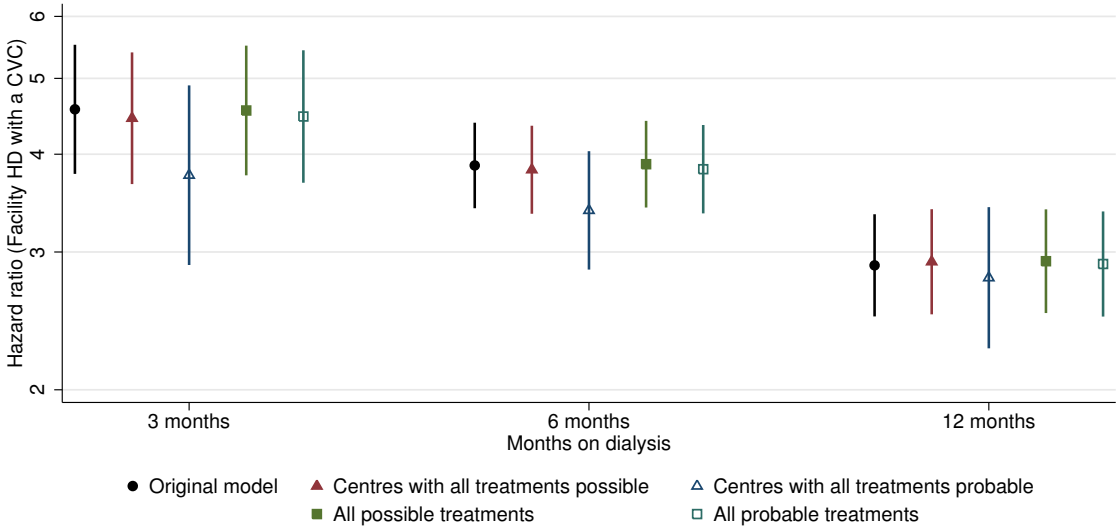
Analyses accounting for clustering by centre C_j, treatments $a \in \mathcal{A}$

Centres with only one treatment possible/probable must be excluded from all analyses.

- Exclude 11 C_j with < 150 periods (545 periods excluded in total)
- Leaves **74 centres, ≈208000 periods**

	Restriction	Included centres	Total no. of periods (1000s)
1	Centres w/ $P(A = a \mid C_j) > 0, \forall a \in \mathcal{A}$	68	192
2	Centres w/ $P(A = a \mid C_j) > 0.05, \forall a \in \mathcal{A}$	34	128
3	Treatments w/ $P(A = a \mid C_j) > 0$	74	208
4	Treatments w/ $P(A = a \mid C_j) > 0.05$	70	207

HRs for Facility HD CVC, accounting for clustering by centre



Outline

1. Motivating example: dialysis and the ANZDATA Registry

2. Clustering by dialysis centre

3. Unmeasured confounding and clustering

Unmeasured confounding and clustering: confounding function

- $D(t) = 1$ if death at time t
- $D_a(t)$: **counterfactual** outcome had this patient received dialysis type a.
- The confounding function:

$$c(a) = \frac{P(D_a(t) = 1 | A(t) = a, V = v)}{\frac{1}{\sum_{a^* \in \mathcal{A} \setminus \{a\}} P(a^*)} \sum_{a^* \in \mathcal{A} \setminus \{a\}} P(a^*) P(D_a(t) = 1 | A(t) = a^*, V = v)},$$

$$P(a^*) = P(A(t) = a^* | V = v).$$

- $c(a)$: **HR of death comparing patients on a to those not on a, had those patients been (contrary to the fact!) on a.**

Unmeasured confounding and clustering: confounding function

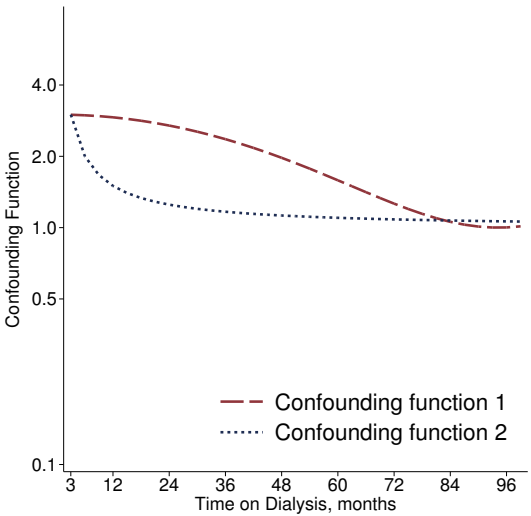
- $D(t) = 1$ if death at time t
- $D_a(t)$: **counterfactual** outcome had this patient received dialysis type a.
- The confounding function:

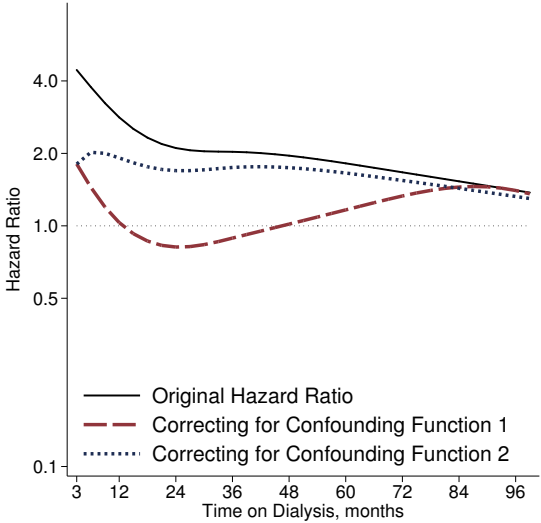
$$c(a) = \frac{P(D_a(t) = 1 | A(t) = a, V = v)}{\frac{1}{\sum_{a^* \in \mathcal{A} \setminus \{a\}} P(a^*)} \sum_{a^* \in \mathcal{A} \setminus \{a\}} P(a^*) P(D_a(t) = 1 | A(t) = a^*, V = v)},$$

$$P(a^*) = P(A(t) = a^* | V = v).$$

- $c(a)$: **HR of death comparing patients on a to those not on a, had those patients been (contrary to the fact!) on a.**
  - $c(a) = 1$: no difference in the risk of death of patients on a and those not on a.
  - **c(Facility HD CVC)** > 1: Facility HD CVC patients have a greater risk of death than those patients on PD/ Home HD/ Facility HD AVF/AVG (had those patients been on Facility HD CVC).
- Possible that the impact of unmeasured confounding differs across dialysis centres.

HRs accounting for unmeasured confounding: Facility HD + CVC





Discussion

- Clustering is not often accounted for in the application of MSMs:
  - The link between clustering and the positivity assumption must be considered!
  - If treatment options are restricted (instead of centres): HRs defined only for those centres in which the treatment is available.
  - There remain questions about the best way to account for clustering.
- Sensitivity of the estimates to the impact of unmeasured confounding should be assessed:
  - Confounding functions are useful in this assessment
- Future research: quantifying the amount of unmeasured confounding accounted for by including cluster effects.

References

B. A. Brumback, M. A. Hérnan, S. J. P. A. Haneuse & J. M. Robins
Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures.
Statistics in Medicine, 23:749–767, 2008.

Z. Fewell, M. A. Hérnan, F. Wolfe, K. Tilling, H. Choi & J. A. C. Sterne.
Controlling for time-dependent confounding using marginal structural models.
The Stata Journal, 4(4): 402-420, 2004.

M. A. Hérnan, B. Brumback & J. M. Robins .
Marginal structural models to estimate the causal effects of Zidovudine on the survival of HIV-positive men.
Epidemiology, 11(5):561-570, 2000.

J. Kasza, K. R. Polkinghorne, M. R. Marshall, S. P. McDonald & R. Wolfe
Clustering and residual confounding in the application of marginal structural models: dialysis modality, vascular access, and mortality.
American Journal of Epidemiology, to appear.