Supporting Information for Analysis of Readmissions Data Taking Account of Competing Risks by Wenbo Wu, Kevin He, Xu Shi, Douglas E. Schaubel, and John D. Kalbfleisch

Web Appendix A Justifying L as a Partial Likelihood

We define the history $\mathbf{H}(k)$ up to time $k = 1, \ldots, \tau$ as

$$\mathbf{H}(k) \coloneqq \left\{ \{Y_{ij}(l), \mathrm{d}N_{ij}^{\mathrm{R}}(l-1), \mathrm{d}N_{ij}^{\mathrm{D}}(l-1), \mathbf{Z}_{ij}(l)\}_{l=1}^{k} : i = 1, \ldots, m, j = 1, \ldots, n_i \right\}.$$

The conditional likelihood of $\mathrm{d}N_{ij}^{\mathrm{R}}(k)$ given $\mathbf{H}(k)$ is

$$f(\mathrm{d}N_{ij}^{\mathrm{R}}(k)|\mathbf{H}(k)) \coloneqq \lambda_{ij}(k)^{Y_{ij}(k)\mathrm{d}N_{ij}^{\mathrm{R}}(k)}[1-\lambda_{ij}(k)]^{Y_{ij}(k)[1-\mathrm{d}N_{ij}^{\mathrm{R}}(k)]}.$$

It follows that

$$\begin{aligned} L &= \prod_{k=1}^{\tau}\prod_{i=1}^{m}\prod_{j=1}^{n_i} \lambda_{ij}(k)^{Y_{ij}(k)\mathrm{d}N_{ij}^{\mathrm{R}}(k)}[1-\lambda_{ij}(k)]^{Y_{ij}(k)[1-\mathrm{d}N_{ij}^{\mathrm{R}}(k)]} \\ &= \prod_{k=1}^{\tau}\left\{ \prod_{i=1}^{m}\prod_{j=1}^{n_i} f(\mathrm{d}N_{ij}^{\mathrm{R}}(k)|\mathbf{H}(k)) \right\}, \end{aligned}$$

in which we have nested σ-algebras, i.e., $\sigma(\mathbf{H}(k)) \subset \sigma(\mathbf{H}(k+1))$, $k = 1, \ldots, \tau - 1$. By Cox (1975) and Wong (1986), L is a partial likelihood based on $\mathrm{d}N_{ij}^{\mathrm{R}}(k)$.

Web Appendix B Justifying $\hat{\boldsymbol{\theta}}$ as a GEE Estimator

To ease notation, for $i = 1, \ldots, m$, $p = 1, \ldots, n^{(i)}$, $l = 1, \ldots, n_{ip}$, and $k = 1, \ldots, \tau$, we define

$$\begin{aligned} \mathbf{U}_{ipl}(k) &\coloneqq [u_{iplk}(1), \ldots, u_{iplk}(m)]^{\top}, \\ \mathbf{W}_{ipl}(k) &\coloneqq [w_{iplk}(1), \ldots, w_{iplk}(\tau)]^{\top}, \end{aligned}$$

where $u_{iplk}(q) = 1(q = j)$ and $w_{iplk}(q) = 1(q = k)$. Then we have

$$h_{iplk} = h(\eta_k + \gamma_i + \mathbf{Z}_{ipl}^\top(k)\boldsymbol{\beta}) = h(\mathbf{X}_{ipl}^\top(k)\boldsymbol{\theta}),$$

in which $\mathbf{X}_{ipl}(k) \coloneqq [\mathbf{U}_{ipl}^\top(k), \mathbf{W}_{ipl}^\top(k), \mathbf{Z}_{ipl}^\top(k)]^\top$. Further, we let

$$\begin{aligned}
\mathbf{D}_{ip} &= [\dot{h}_{ip11}\mathbf{X}_{ip1}(1), \ldots, \dot{h}_{ip1\tau}\mathbf{X}_{ip1}(\tau), \ldots, \dot{h}_{ipn_{ip}1}\mathbf{X}_{ipn_{ip}}(1), \ldots, \dot{h}_{ipn_{ip}\tau}\mathbf{X}_{ipn_{ip}}(\tau)]^\top, \\
\mathrm{d}\mathbf{N}_{ip}^{\mathrm{R}} &\coloneqq [\mathrm{d}N_{ip1}^{\mathrm{R}}(1), \ldots, \mathrm{d}N_{ip1}^{\mathrm{R}}(\tau), \ldots, \mathrm{d}N_{ipn_{ip}}^{\mathrm{R}}(1), \ldots, \mathrm{d}N_{ipn_{ip}}^{\mathrm{R}}(\tau)]^\top, \\
\mathbf{Y}_{ip} &\coloneqq \mathrm{diag}(Y_{ip1}(1), \ldots, Y_{ip1}(\tau), \ldots, Y_{ipn_{ip}}(1), \ldots, Y_{ipn_{ip}}(\tau)), \\
\mathbf{h}_{ip} &\coloneqq [h_{ip11}, \ldots, h_{ip1\tau}, \ldots, h_{ipn_{ip}1}, \ldots, h_{ipn_{ip}\tau}]^\top, \\
\mathbf{V}_{ip} &\coloneqq \mathrm{diag}(v(h_{ip11}), \ldots, v(h_{ip1\tau}), \ldots, v(h_{ipn_{ip}1}), \ldots, v(h_{ipn_{ip}\tau})), \\
\mathbf{b}_{ip} &\coloneqq [b(h_{ip11}), \ldots, b(h_{ip1\tau}), \ldots, b(h_{ipn_{ip}1}), \ldots, b(h_{ipn_{ip}\tau})]^\top,
\end{aligned}$$

where h_{iplk} is a shorthand of $h(\eta_k + \gamma_i + \mathbf{Z}_{ij}^\top(k)\boldsymbol{\beta})$, and b and v are known functions given by

$$b(h) \coloneqq \sqrt{h(1-h)}, \quad v(h) \coloneqq \frac{\dot{h}^2}{h(1-h)},$$

with $\dot{h}$ representing the first-order derivative of h.

Following the GEE framework, we have a system of unbiased estimating equations

$$\sum_{i=1}^{m}\sum_{p=1}^{n^{(i)}} \mathbf{D}_{ip}^\top \mathbf{V}_{ip}^{-1}(\mathrm{d}\mathbf{N}_{ip}^{\mathrm{R}} - \mathbf{Y}_{ip}\mathbf{h}_{ip}) = \mathbf{0}, \tag{1}$$

in which an independent working correlation matrix has been assumed. With some algebra, system (1) reduces to

$$\phi^{-1}\sum_{i=1}^{m}\sum_{p=1}^{n^{(i)}}\sum_{l=1}^{n_{ip}}\sum_{k=1}^{\tau}\mathcal{U}_{iplk}\mathbf{X}_{ipl}(k) = \mathbf{0},$$

which is easily seen to be equivalent to $\mathcal{U}(\boldsymbol{\theta}) = \mathbf{0}$. This implies that the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ is a solution to (1), i.e., $\hat{\boldsymbol{\theta}}$ is a GEE estimator under the working independence assumption.

Web Appendix C Assumptions on the Stabilized Robust Variance Estimator

As in a marginal model, we make three assumptions on the stabilized robust variance estimator.

Assumption 1. The conditional mean of $\mathrm{d}N_{ipl}^{\mathrm{R}_*}(k)$ (i.e., cause-specific hazard of readmission)

depends on the covariates $\mathbf{Z}_{ipl}(k)$ via a known function, that is,

$$\mathrm{E}(\mathrm{d}N^{\mathrm{R}_*}_{ipl}(k)|T_{ipl} \geq k, \mathbf{Z}_{ipl}(k)) = \lambda_{ipl}(k) = h_{iplk}. \tag{2}$$

Assumption 2. The conditional variance of $\mathrm{d}N^{\mathrm{R}_*}_{ipl}(k)$ depends on the conditional mean h_{iplk} in (2) according to

$$\mathrm{Var}(\mathrm{d}N^{\mathrm{R}_*}_{ipl}(k)|T_{ipl} \geq k, \mathbf{Z}_{ipl}(k)) = v(h_{iplk})\phi,$$

where $\phi > 0$ is an unknown scale parameter.

Assumption 3. For any $n \in \{1, \ldots, \bar{n}\}$, all patients who have a τnth-order leading principal submatrix of the correlation matrix of $\mathrm{d}\mathbf{N}^{\mathrm{R}}_{ip}$ share a common structure for that submatrix.

Compared with Pan (2001), Assumption 3 extends the notion of common correlation structure to patient-level clustering with unequal discharge counts. An underlying justification is that for those patients who have experienced at least n discharges, their levels of exposure to additional UHRs should be similar, regardless of the different discharge counts they would eventually have. This assumption allows inter-patient discharge pooling to be independent of patient-specific discharge count, which, implicitly, is a random variable for every patient.

Web Appendix D Alternative Tests

A generalized Wald test statistic is given by

$$T_i^{\mathrm{RW}} := \frac{\hat{\gamma}_i - \hat{\gamma}_{\mathrm{M}}}{\sqrt{\widehat{\Sigma}_{\gamma_i}}},$$

which has the same asymptotic distribution as T_i^{RS}.

If the competing risk model is correctly specified, a model-based score test statistic under the null hypothesis $H_{0i} : \gamma_i = \gamma_{\mathrm{M}}$ can be expressed as

$$T_i^{\mathrm{MS}} \frac{\mathcal{U}_i(\tilde{\boldsymbol{\gamma}}_i)}{\sqrt{\mathcal{I}_i(\tilde{\boldsymbol{\gamma}}_i)}},$$

where $\mathcal{I}_i(\tilde{\boldsymbol{\gamma}}_i)$ is the ith diagonal element of the information matrix $\mathcal{I}(\tilde{\boldsymbol{\gamma}}_i)$ evaluated at $\tilde{\boldsymbol{\theta}}_i$. Likewise, a model-based Wald test statistic is given by

$$T_i^{\mathrm{MW}} := \sqrt{\mathcal{I}_i(\hat{\boldsymbol{\gamma}})}(\hat{\gamma}_i - \hat{\gamma}_{\mathrm{M}}),$$

where $\mathcal{I}_i(\hat{\boldsymbol{\gamma}})$ is the ith diagonal element of $\mathcal{I}(\hat{\boldsymbol{\gamma}})$ evaluated at $\hat{\boldsymbol{\theta}}$. Under certain regularity conditions, T_i^{MS} and T_i^{MW} also have an asymptotic standard normal distribution.

Web Appendix E Simulation Details

As acknowledged in the literature (Prentice, 1988; Pan, 2001; Diggle et al., 2002), with a logit link and clustered binary responses, it is unlikely to build a model in which subjects with the same number of repeated measurements share a common correlation matrix structure. Since our inference framework bears a resemblance to a marginal model, the previous observation carries over to our simulation setting. Nonetheless, Assumptions 1 to 3 in Appendix D can be approximately satisfied by assuming the following misspecified model with random effect:

$$\mathrm{E}(\mathrm{d}N^{\mathrm{R}_*}_{ipl}(k)|T_{ipl} \geq k, \mathbf{Z}_{ipl}(k), \varepsilon_{ipl}) = h(\eta_k + \gamma_i + \mathbf{Z}^\top_{ipl}(k)\boldsymbol{\beta} + \varepsilon_{ipl}),$$

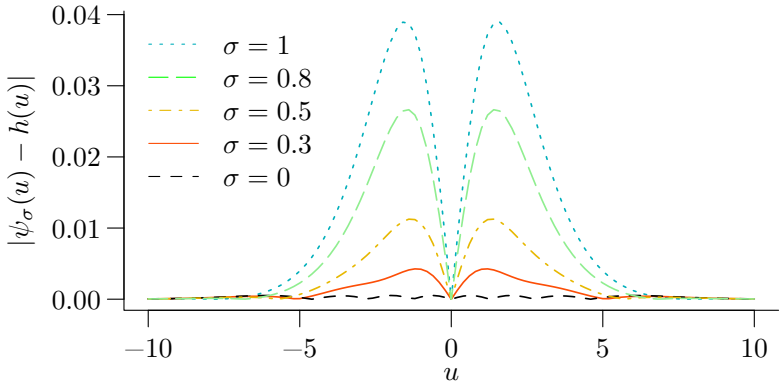
where $\boldsymbol{\varepsilon}_{ip} = [\varepsilon_{ip1}, \ldots, \varepsilon_{ipn_{ip}}]^\top \sim \mathcal{N}_{n_{ip}}(\mathbf{0}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ being an exchangeable covariance matrix with marginal variance σ^2 and correlation ρ. When h is a standard logistic function (corresponding to the logit link), we have

$$\mathrm{E}(\mathrm{d}N^{\mathrm{R}_*}_{ipl}(k)|T_{ipl} \geq k, \mathbf{Z}_{ipl}(k)) \approx \psi_\sigma(\eta_k + \gamma_i + \mathbf{Z}^\top_{ipl}(k)\boldsymbol{\beta}),$$

where

$$\psi_\sigma(u) \coloneqq \kappa \cdot \Phi\left(\frac{u}{\sqrt{\xi_1 + \sigma^2}}\right) + (1-\kappa) \cdot \Phi\left(\frac{u}{\sqrt{\xi_2 + \sigma^2}}\right),$$

with $\kappa = 0.4353$, $\xi_1 = 2.2967^2$, and $\xi_2 = 1.3017^2$, and the last so-called two-probit approximation is due to Demidenko (2013, §7.7.2, pp.406). Web Figure 1 displays absolute errors of two-probit approximation with respect to the standard logistic function under different values of σ. In particular, $\sigma = 0.3$ strikes a balance between introducing correlation and approximating the logistic function, with a maximum absolute error of 0.0043.



Web Figure 1: Absolute errors of two-probit approximation relative to standard logistic function under different values of σ.

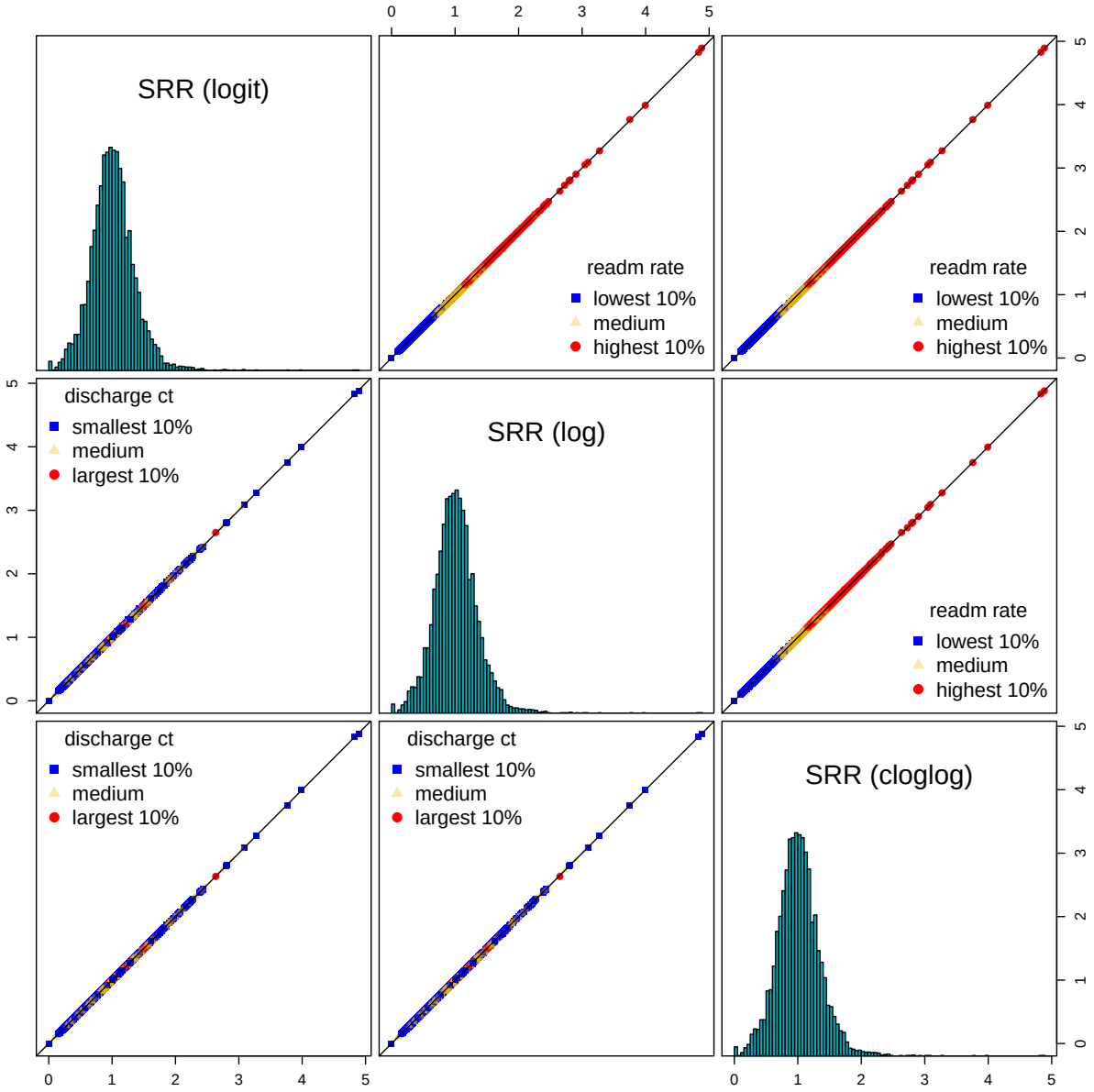
Web Appendix F Application Details

The readmissions data are derived from an extensive national ESRD patient database garnered from multiple data systems operated by federal public health agencies in the United States. Significant sources include the Renal Management Information System, CROWN-Web facility-reported clinical and administrative data (extracted from CMS-2728 Medical Evidence Form, CMS-2746 Death Notification Form, and CMS-2744 Annual Facility Survey Form), Medicare Enrollment Database, Medicare claims data from Standard Analytic Files, transplant data from the Scientific Registry of Transplant Recipients, nursing home Minimum Data Set, provider survey and certification data from Quality Improvement and Evaluation System Business Intelligence Center, and the Dialysis Facility Compare.

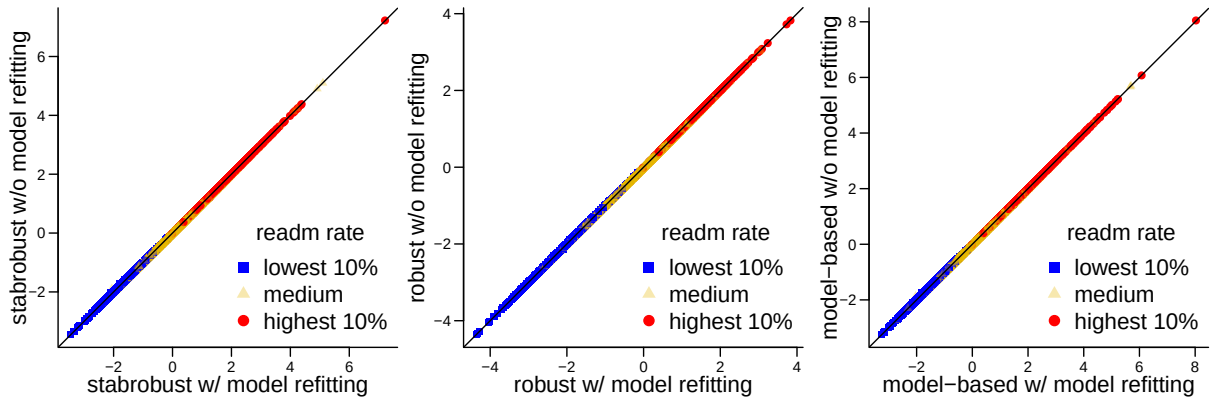
Applying exclusion criteria (such as patients with a primary diagnosis of cancer, mental health, or rehabilitation, discharged against medical advice, or hospitalized at Prospective Payment System-exempt cancer hospitals) to qualifying discharges taking place between January 1 and December 31, 2018, there were 541,769 discharges (257,860 patients) included in analysis, each with at least 11 discharges. The 6,937 facilities had a wide variety of discharge counts from 11 (52 facilities) to 842 (1 facility) with a mean of 78.10, readmission counts from 0 (19) to 264 (1) with a mean of 20.58, and competing event counts from 0 (707) to 76 (1) with a mean of 3.72. Among them, 10% had at most 24 discharges or 5 readmissions, and 26.35% had at most 1 competing risk. Patient-specific discharge counts spanned from 1 (143,704 patients) to 12 (139 patients), with 75% of patients having at most 2 discharges. The observed 30-day facility-specific readmission rates ranged from 0% (19 facilities) to 73.33% (1 facility), with an overall readmission rate of 26.35%; the observed 30-day facility-specific rates of competing risks varied from 0% (707) to 23.08% (1), with an overall rate of 4.76%. We consider 74 predictors, including age, sex, body mass index, years on dialysis, status of Medicare Advantage Plans at discharge, length (days) of index hospitalization, diabetic status, past-year nursing home status at discharge, past-year prevalent comorbidities and high-risk conditions at discharge (using Agency for Healthcare Research and Quality Clinical Classifications Software ICD-10 diagnosis categories), and days (4 to 30) to the first event after discharge.

Web Appendix G Supplementary Figures

Web Figure 2 implies that different link functions give rise to similar distribution of SRRs in our application setting. Web Figure 3 demonstrates that using estimates from the constrained model refitting procedures yield similar score test statistics as those without model refitting, regardless of which variance estimator is considered.



Web Figure 2: SRR from competing risk models with different link functions. Histograms are in the diagonal panels. Facilities are stratified by readmission rate or discharge count. Dashed lines represent 2.5% and 97.5% quantiles of the standard normal distribution. 45-degree lines in solid black.



Web Figure 3: Score test statistics with versus without constrained model refitting using different variance estimators. "stabrobust", "robust" and "model" correspond to test statistics with the stabilized robust, classical robust and model-based variance estimators, respectively. Facilities are stratified by readmission rate. 45-degree lines in solid black.

References

Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2):269–276.

Demidenko, E. (2013). *Mixed Models: Theory and Applications with R*. John Wiley & Sons.

Diggle, P. J., Heagerty, P., Liang, K.-Y., and Zeger, S. (2002). *Analysis of Longitudinal Data*. Oxford University Press.

Pan, W. (2001). On the robust variance estimator in generalised estimating equations. *Biometrika*, 88(3):901–906.

Prentice, R. L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics*, 44(4):1033–1048.

Wong, W. H. (1986). Theory of partial likelihood. *Annals of Statistics*, 14(1):88–123.