

On Assessing the Quality of Health Care Providers Using Time-Varying Patient Outcomes

Sebastien Haneuse,^a José Zubizarreta,^b and Sharon-Lise T Normand^{a,b}

^aDepartment of Biostatistics and ^bDepartment of Health Care Policy
Harvard T.H. Chan School of Public Health and Harvard Medical School
Boston, MA, U.S.A.

Estes et al [1] provide an interesting and intriguing approach to assessment of dialysis facilities for time-varying patient outcomes. They develop a method to monitor patient outcomes continuously over time for a population of patients with end-stage renal disease. It is our belief that statistics has not played a sufficiently prominent role in the development of health care policy, and we congratulate the authors in their efforts to push the field forward. From a technical standpoint, we find particularly interesting the extension of multilevel varying coefficient models (MVCMs) for time-varying hospital profiling. However, we also have concerns about the practical usefulness of this work and its potential to influence policy. We discuss the policy question as well as the interpretation and robustness of conclusions. We begin with some background context for our discussion.

Background

The general rationale for estimating hospital readmission rates are their relationships to patient morbidity, quality of life, and medical costs. The Centers for Medicare and Medicaid Services (CMS), a branch within the U.S. government responsible for managing federal health care programs, estimate hospital-specific readmission measures for several disease-based cohorts, such as patients with heart failure or pneumonia, as well as surgical cohorts, such as patients undergoing cardiac or orthopedic surgery. Among the many metrics used for evaluating hospital quality of care, hospital-specific readmission rates are particularly prominent because they are used to modify hospital payments – in fiscal year 2017, for example, CMS withheld more than \$500 million in readmission penalties from about 2500 hospitals [2].

When using readmission rates as a means to evaluate hospital performance, one of two questions is typically asked. The first asks,

How does the readmission rate of a given hospital compare to that of other hospitals?

This question can be addressed using direct standardization with respect to some population of patients, typically the entire patient population across all hospitals [7]. The second question asks,

How would the patients treated at a given hospital have fared on average across all hospitals?

This question can be addressed via indirect standardization to the specific case-mix treated at the given hospital [3]. Central to both questions is the understanding that the hospital is responsible for post-discharge care, at least during some time window, typically 30 days. CMS examines all live discharges over a 12-month period for a hospital and estimates the number of unplanned hospital readmissions within 30-days of hospital discharge relative to the expected number of unplanned readmissions for the hospital. The estimates are indirectly standardized and, as indicated, CMS withholds some percentage of the hospital payments for those hospitals having a higher than expected number of 30-day readmissions. The public can see these and other hospital-specific estimates of quality on the CMS website *Hospital Compare*.

While similar in many respects, the readmission metric used in *Dialysis Facility Compare* is different than that used in *Hospital Compare* in that the attribution of responsibility for post-discharge care to the dialysis facility is without regard as to where the hospitalization took place [8]. Specifically, the metric reported by *Dialysis Facility Compare* is the ratio of the number of Medicare-covered index hospital discharges from any acute care hospital that resulted in an unplanned readmission within 4-30 days of discharge observed for dialysis patients attributed to a particular dialysis facility relative to the number of unplanned readmissions that would be expected at hospitals discharging the dialysis center's patients. Denoting this metric as the *standardized readmission ratio* (SRR), case-mix differences across facilities are adjusted using of patient- and discharge-specific characteristics, including sex, cause of end stage renal disease, comorbidity history, age at discharge, time on dialysis at discharge and length of stay of the index hospitalization.

What is the Policy-Relevant Question Addressed by Estes et al?

While the SRR was not proposed by Estes and colleagues, they do propose the *standardized dynamic readmission ratio* SDRR(t) as an extension, with t indexing time since initiation of dialysis. In seeking a quality metric that acknowledges variation over time, however, a number of important changes to the core methodology of the SRR are made. First, case-mix adjustment in their proposed framework is restricted to characteristics measured at the time of initiation of dialysis; that is, in contrast to both *Dialysis Facility Compare* and *Hospital Compare*, no information on patients that is specific to the hospitalization or discharge is used. The authors do acknowledge this, citing reasonable concerns regarding the role of time-dependent confounding. However, while restricting case-mix adjustment in this way may account for variation across dialysis facilities close to the time of dialysis initiation, it seems likely that residual confounding at later time points will be inevitable. Second, in developing a time-varying measure, the authors were required to acknowledge the role of mortality as a competing risk. The specific approach taken was to

adopt a model specification that conditions on a patient being alive. Thus, at t days from dialysis initiation, the $\text{SDRR}(t)$ is an estimate of the number of readmissions for dialysis patients in the facility who are alive relative to the expected number of readmissions for an average facility having the same types and numbers of patients alive at time t . Originally developed as a means to account for death in longitudinal data analysis, a potential problem with the use of this so-called *partly conditional* approach is that there may be important variation across dialysis facilities in ‘quality’ with respect to mortality. This, in turn, could lead to an erroneous conclusion that a facility has good performance with respect to readmission that is purely due to poor performance with respect to mortality [5].

Beyond these methodological differences, and how they can be explained and justified to relevant stakeholders, we believe that the interpretation of the proposed $\text{SDRR}(t)$ metric suffers as a measure of quality. Moreover, we find it difficult to attribute responsibility of post-discharge care following a hospitalization to the dialysis center rather than to the hospital where the hospitalization took place because: (i) dialysis centers are specialized outpatient facilities and, as such, patients who experience a clinical event that warrants hospitalization (e.g. heart failure or pneumonia), will not be hospitalized at the center itself; (ii) patients undergoing dialysis at a given facility may (and will likely) be hospitalized at different hospitals; and, (iii) shover a 3-year window that is being considered, the set of hospitals ‘contributing’ hospitalizations to a dialysis facilities metric may vary. We acknowledge that the first two of these points apply to the SRR measure currently reported by *Dialysis Facility Compare*. Nevertheless, they apply to $\text{SDRR}(t)$ with the third (additional) point arising because of the time-varying nature of the metric. Finally, setting aside these issues, it is unclear how one should relate the output from the approach (i.e. the curves presented in Figure 1 of Estes et al [1]) to some notion of ‘quality’. Consider, for example, the hospital-specific curve in the ‘Small: Not different’ panel that decreases from $\text{SDRR}(0) \approx 3.0$ to $\text{SDRR}(300) \approx 0.9$, back up to $\text{SDRR}(800) \approx 2.0$ and finally to $\text{SDRR}(1000) \approx 0.5$. From this, one is to infer that the facility’s performance was poor immediately after dialysis started, then became good, then became poor again, and finally was excellent. While we certainly understand and appreciate the general benefits of flexibility in model-specification, in the context considered here, how can such variation be explained by hospital characteristics, dialysis facility characteristics, or how a hospital or dialysis facility is administered.

We also wonder about the robustness of the proposed approach to limited overlap in patients covariate distributions across hospitals, across time. If there is limited overlap in patients covariate distributions, then the results can be very sensitive to misspecification of the outcome regression models. As argued by Silber et al. [6], a regression model can say what it has seen, and in hospitals where there are only certain types of patients, it will extrapolate or fabricate results for other types of patients that are absent. While this issue manifests in the time-invariant setting, it is exacerbated by the added time dimension in profiling as the available data are spread more thinly. We suggest that methods like those proposed by Estes et al. be accompanied by overlap diagnostics to better understand to what extent the available data supports the dialysis center comparisons.

Concluding Remarks

We have very much appreciated the opportunity to discuss this article and thank the authors for their work. Collectively, in our view, the issues we raise speak, in one way or another, to the fundamental issue of what is the policy-relevant question being addressed? If the ultimate goal is to guide the development of quality-improvement interventions/programs, policy-makers, dialysis facility administrators and other stakeholders will need to be convinced that metrics used are meaningful and fair. Given this, is the proposed time-dynamic provider assessment useful? While we agree that evaluating medical providers on the basis of recurrent or time-varying patient outcomes would be useful, particularly for patients with chronic conditions, a strong case needs to be made that links the quality of care delivered by the provider to the subsequent patient outcomes. Even in the, arguably, more straightforward setting of hospitalizations due to heart failure, it is challenging to convince hospitals that they are responsible for unplanned hospital readmissions within 30-days of a hospital discharge. We have difficulty seeing the link of unplanned hospital readmissions over time to dialysis facilities. We acknowledge that we have raised numerous issues without any concrete solutions but appreciate that Estes et al have initiated an important conversation.

Acknowledgments

S-L Normand's effort was supported, in part, by Grant R01-GM111339 from the National Institute of General Medical Sciences, U.S. National Institute of Health.

References

1. Estes JP, Nguyen DV, Chen Y, Dalrymple LS, Rhee CM, Kalantar-Sazeh K, Senturk D (2018). Time-dynamic profiling with application to hospital readmission among patients on dialysis. *Biometrics* XXX.
2. Henry Kaiser Family Foundation (2017). Aiming for fewer hospital u-turns: the Medicare Hospital Readmission Reduction Program, March 10 (Issue Brief).
3. Normand SLT, Glickman ME, Gatsonis CA. (1997). Statistical methods for profiling providers of medical care: issues and applications. *Journal of the American Statistical Association*, 92(439), 803-814.
4. Normand SLT, Shahian DM. (2007). Statistical and clinical aspects of hospital outcomes profiling. *Statistical Science*, 206-226.
5. Lee KH, Dominici F, Schrag D, Haneuse S (2016). Hierarchical models for semicompeting risks data with application to quality of end-of-life care for pancreatic cancer. *Journal of the American Statistical Association*, 111(515), 1075-1095.
6. Silber JH, Rosenbaum PR, Ross RN, Ludwig JM, Wang W, Niknam BA et al (2016). Indirect standardization matching: assessing specific advantage and risk synergy. *Health services research*, 51(6), 2330-2357.

7. Vawecyj N, Goetghebeur E, Eriksson M, Vansteelandt S (2014). On shrinkage and model extrapolation in the evaluation of clinical center performance. *Biostatistics*, 15:651-664.
8. University of Michigan Kidney Epidemiology and Cost Center. Dialysis Facility Compare – Standardized Readmission Ratio <https://dialysisdata.org/sites/default/files/content/Methodology/SRR%20MIF.pdf>