The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: Jump to Page 1

Support RAND

Browse Reports & Bookstore
Make a charitable contribution

For More Information

Visit RAND at www.rand.org
Explore RAND Health
View document details

Limited Electronic Distribution Rights
This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see RAND Permissions.
This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.
Provider-Level Risk-Adjusted Quality Measurement for Inpatient Rehabilitation Facilities

Andrew W. Dick, Peter J. Huckfeldt, Hangsheng Liu, Hao Yu, Ateev Mehrotra, Susan L. Lovejoy, J. Scott Ashwood

Sponsored by the Medicare Payment Advisory Commission
The research described in this report was sponsored by the Medicare Payment Advisory Commission and was conducted within RAND Health, a unit of the RAND Corporation.
SUMMARY

Quality metrics play an increasingly important role in the evaluation and reimbursement of post-acute providers, and inpatient rehabilitation facilities (IRFs) in particular. The Medicare Payment Advisory Commission (MedPAC) uses aggregate trends in quality measures, such as functional gain, to assess the adequacy of payments to providers. In addition, the Patient Protection and Affordable Care Act (ACA) requires IRFs to publicly report quality data or be penalized in annual payment updates. Eventually, quality data could be used in a value-based purchasing scheme for IRFs as is being developed for other post-acute settings under ACA.

In this report, we describe our development of risk-adjusted quality metrics for IRFs, focusing on five patient outcomes: (1) functional gain, (2) discharge to the community, (3) 30-day readmission to acute care, (4) 30-day readmission to a skilled nursing facility (SNF), and (5) discharge directly to acute care. We start by describing the development of our model, including our choice of independent variables, the selection of our model specification, and the identification of a minimum patient volume to support a quality metric. Next, we present results from our risk-adjustment estimation. We identify changes in model parameter estimates across years in the sample period and decompose average patient outcome trends into changes stemming from the case-mix composition of patients seen in IRFs and those due to changes in expected outcomes given the observable case mix. Finally, we examine the persistence of quality estimates and differential trends in quality by IRF provider characteristics.

For each outcome, we specified a model that contained individual socioeconomic and demographic characteristics, comorbid condition indicators, Impairment Group Code (IGC) indicators, and age-by-sex interactions. We estimated nine alternative specifications for each model and chose the model with the best fit. In each case, the best model was either the full model or the model that dropped the age-by-sex interactions. Model fit was good, but there remained a substantial amount of unexplained outcome variation by IRF.
We considered two main modeling approaches for the IRF risk adjustment: (1) random-effects or hierarchical models and (2) fixed-effects models in which each IRF’s contribution to quality is determined by a fixed indicator in the model. For each of the outcomes models, we found strong evidence to reject the random-effects assumption, indicating that quality measures based on random-effects estimates would be biased. We therefore adopted the fixed-effects approach. In the case of this application, even though the number of IRFs was large (1,400), the very large data set (2.5 million observations) made it feasible to estimate fixed-effects models. One notable advantage of the fixed-effects approach is that the models generate an empirical representation of the distribution of quality across IRFs. Unlike random-effects models, which assume a distribution for IRF quality, the fixed-effects approach imposes no structure. In addition, by manipulating the models so that each IRF’s estimate represents its difference from the average-quality IRF, the quality metrics are relatively easy to manipulate and interpret.

Although our study sample contained approximately 2.5 million observations, not every IRF had large patient volume. Thus, we developed criteria for determining if the quality estimate for an IRF was too uncertain or unreliable to merit reporting. Starting with the assumption that the estimated uncertainty (standard error or confidence interval of the quality measure) is the best means for determining the measure’s reliability, we examined the relationship between IRF volume and the quality measure’s uncertainty, and we examined various criteria that used only IRF volume to determine the omission of IRFs. We found that the use of a volume standard resulted in either the omission of too many IRFs (many of which had reliable estimates of quality) or the inclusion of IRFs that had very uncertain quality estimates. We therefore recommend the use of a combined standard that requires a minimum patient volume (we used n = 30) to eliminate very-small volume IRFs and a maximum uncertainty level (confidence interval > one standard deviation of the outcome rate across institutions) in order to exclude IRFs for which the quality estimates were not precise. We found that these criteria together resulted in the exclusion of a very small number of IRFs.

We examined model stability by comparing model estimates based on data from early in the study period with model estimates based on data from late in the study period. In
every case but one (30-day readmission to SNF comparing 2006–7 estimates with 2008–9 estimates), we found evidence that the model parameter estimates changed over time. This implies that caution should be used when making model estimates using data pooled over several years. It also suggests that risk-adjustment models will need to be recalibrated regularly. As a result, we estimated year-specific risk-adjustment models to generate year-specific quality indicators for each IRF. These indicators can be compared either within each year, to reveal any IRF’s quality relative to other IRFs or to the average IRF, or over time to reveal trends in quality.

We also used the models to examine trends in quality during the study period. A comparison of raw and adjusted outcome rates shows the contribution of real quality changes and case-mix changes, revealing that real improvements in quality have been masked by increasing severity in case mix during each year of the study and for each of the outcomes considered. Among the key findings are the following:

1. **The raw outcome rates for four of the five quality measures worsened over the study period.** Functional Independence Measure (FIM) gain improved by about two points, but discharge to the community (18 percent to 11 percent), 30-day readmission to acute care (10.5 percent to 11.5 percent), 30-day readmission to SNF (3.1 percent to 3.7 percent), and discharge directly to acute care (8.6 percent to 10.25 percent) all worsened.

2. **The declining raw rates were caused by worsening case mix during the period.** After adjusting for case mix, we found that quality improved on every metric; in the case of FIM gain, quality improved by more than the improvement in the raw rate.

3. **Quality at individual IRFs persisted over the study period.** We identified two cohorts of IRFs using 2004 quality estimates: the lowest-quartile performers and the highest-quartile performers. We then examined the average quality of those cohorts in each of the study years, and we found that the 2004 high-performing IRFs continued to be high performers throughout, that the 2004 low performers continued to be low performers, and that the gap in quality did not narrow substantially.
Our work indicates that the existing data provide ample information to estimate risk-adjustment models for the purpose of examining and reporting IRF quality. Furthermore, our findings show that, because of the substantial worsening of case mix, the use of high-quality risk-adjustment models as the basis for quality reporting is essential for revealing overall quality trends in the market. Finally, we find that the risk-adjustment parameter estimates change over time, indicating that the models should be recalibrated regularly and, if feasible, year-specific models should be used.