2.5 MODEL SPECIFICATION AND VALIDATION

To develop the Hybrid eHWR, we tested and compared three different risk-adjustment approaches using the CCDE and the original HWR measure. All strategies were variations on the basic HWR structure which models the outcome for each of 5 specialty cohorts. For each strategy we made analogous modifications to each of the 5 models.

For model development we used logistic regression models, with outcome Y_i for the i^{th} patient equal to 1 if the patient was readmitted within 30 days of discharge and 0 otherwise. In contrast with the final models described below for calculating the measure, logistic regression models are substantially less computationally intensive, and development using models with fully specified error structures would have taken prohibitively long. Also, by using logistic regression models that did not account for hospital effects, we were able to assess risk factors and model performance without reference to the variation in performance across hospitals. We developed separate logistic regression models of unplanned readmission using the three separate risk-adjustment strategies and the original HWR measure approach listed in Section 2.4. We compared the discrimination for each specialty cohort across the four different models. We selected the best-performing alternative model based on discrimination in terms of the C-statistic. The two alternative models with lower discrimination were discarded. We then continued measure development and testing only for the best-performing model containing the CCDE.

After identifying the best alternative approach using the ordinary logistic regression patient-level model, we used hierarchical logistic regression to model the log-odds of readmission for each of the five cohorts to account for patient clustering within hospitals.[8] This is also consistent with the original fully specified HWR models. We then compared the results of this best approach with the results from original HWR measure approach. Readmission within 30 days was modeled as a function of patient- level demographics, clinical characteristics, comorbidities, and a random hospital-level intercept. This model specification accounts for within-hospital correlation of the observed outcomes and models the assumption that underlying differences in quality among the health care facilities being evaluated lead to systematic differences in outcomes. We estimated a separate hierarchical logistic regression model for each specialty cohort.

Specifically, for a given specialty cohort, we estimated a hierarchical logistic regression model as follows. Let Y_{ij} denote the outcome (equal to 1 if patient *i* is readmitted within 30 days, zero otherwise) for patient i at hospital *j*; $\mathbf{Z_{ij}}$ denotes a set of risk factors. We assume the outcome is related linearly to the covariates via a logit function with dispersion:

$$\text{logit}(\text{Prob}(Y_{ij} = 1)) = \alpha_j + \boldsymbol{\theta}^{*}\mathbf{Z_{ij}} + \varepsilon_{ii} \qquad (1)$$

$$\alpha_j = \mu + \omega_j \; ; \; \omega_j \sim N(0,\tau^2)$$

where $\mathbf{Z_{ij}} = (Z_1, Z_2, \ldots Z_k)$ is a set of *k* patient-level covariates. α_j represents the hospital specific intercept; μ is the adjusted average outcome over all hospitals; and τ^2 is the between hospital variance component and $\varepsilon \sim N(0,\sigma^2)$ captures any over- or under-dispersion. The hierarchical logistic regression model for each cohort was estimated using the SAS software system (GLIMMIX procedure).

Hospital performance assessment

The previous section describes how the models for each specialty cohort are specified and estimated, using a separate hierarchical logistic regression model for that cohort. Each model is then used to calculate a standardized risk ratio (SRR) for each hospital which contributes index admissions to that model. These SRRs, weighted by volume, are then pooled for each hospital to create a composite hospital-wide SRR.

We used the results of each hierarchical logistic regression model to calculate the predicted number of readmissions and the expected number of readmissions at each hospital. The predicted number of readmissions in each cohort was calculated, using the corresponding hierarchical logistic regression model, as the sum of the predicted probability of readmission for each patient, including the hospital- specific (random) effect. The expected number of readmissions in each cohort for each hospital was similarly calculated as the sum of the predicted probability of readmission for each patient, ignoring the hospital specific (random) effect. Using the notation of the previous section, the model specific risk standardized readmission ratio is calculated as follows. To calculate the predicted number of admissions $pred_{Cj}$ for index admissions in cohort C=1,...,5 at hospital *j*, we used

$$pred_{Cj} = \Sigma logit^{-1}(\alpha_j + \boldsymbol{\beta}^{*}\mathbf{Z}_{ij}) \quad (2)$$

where the sum is over all m_{Cj} index admissions in cohort C with index admissions at hospital *j*. To calculate the expected number exp_{Cj} we used

$$exp_{Cj} = \Sigma logit^{-1}(\mu + \boldsymbol{\beta}^{*}\mathbf{Z}_{ij}) \quad (3)$$

Then, as a measure of excess or reduced readmissions among index admissions in cohort C at hospital *j*, we calculated the standardized risk ratio SRR_{Cj} as

$$SRR_{Cj} = pred_{Cj}/exp_{Cj} \quad (4)$$

Risk-standardized hospital-wide 30-day readmission rate

To report a single readmission score, the separate specialty cohort SRRs were combined into a single value. We created a single score as follows.

For a given hospital, j, which has patients in some subset of cohorts $C \subseteq {}^{1}$, calculate the SRR as described above for each specialty cohort for which the hospital discharged patients. If the hospital does not have index admissions in a given cohort c, then $m_{cj} = 0$ and we take $SRR_{cj} = 1$. Then, calculate the volume-weighted logarithmic mean:

$$SRR_j = \exp((\Sigma\, m_{cj} \log(SRR_{cj})) / \Sigma m_{cj}) \quad (5)$$

where the sums are over all specialty cohorts; note that if a hospital does not have index admissions in a given cohort ($m_{cj} = 0$) then that cohort contributes nothing to the overall score SRR_j. **This value, SRR_j, is the hospital-wide standardized risk ratio** for hospital *j*. To aid interpretation, this ratio is then multiplied by the overall raw readmission rate for all index admissions in all cohorts for the 21 KPNC hospitals, to produce **the risk-standardized hospital-wide readmission rate ($RSRR_j$)**.

$$RSRR_j = SRR_j * \bar{Y} \quad (6)$$

Model Performance Assessment

We completed hierarchical modeling and calculated measure results for the original HWR model and for the best-performing model containing the CCDE, which we have referred to as the Hybrid eHWR. Assessment of the Hybrid eHWR performance included model calibration (to assess over-fitting), discrimination in terms of predictive ability (the range of observed readmission rates across deciles of predicted rates), and distribution of model residuals. These analyses were done in the development, validation, and testing (2012) samples. We also calculated the model estimates as well as the coefficients and 95% confidence intervals for risk-adjustment variables for the best-performing model in the development and validation samples.

2.6 MEASURE TESTING

To assess the overall internal consistency of the specialty cohort SRRs and appropriateness of combining the SRRs into a composite score, we calculated Cronbach's coefficient α. This coefficient reflects the proportion of total variance in the summated scale composite score that is accounted for by a common source among the condition measures. Theoretically, α varies from 0 to 1 and higher values of α are more desirable.

To determine the extent to which the assessments of a hospital using different but randomly selected subsets of patients produces similar measures of hospital performance, we calculated the RSRR from the Hybrid eHWR using each half of the split-sample 2010-2011 data (the development and validation samples). Thus, we obtain two RSRRs for each hospital, using an entirely distinct set of patients from the same time period. To the extent that the calculated measures of these two subsets agree, we have evidence that the measure is assessing an attribute of the hospital, not of the patients. As a metric of agreement we calculated the intra-class correlation as defined by ICC (2,1) by Shrout and Fleiss (1979).[9,10] For the hospital event rate based on the patient binomial outcomes like readmission (Yes/No), an ICC value of 0-0.2 indicates poor agreement; 0.3-0.4 indicates fair agreement; 0.5-0.6 indicates moderate agreement; 0.7-0.8 indicates strong agreement; and >0.8 indicates almost perfect agreement.[10]

We considered all measure testing as preliminary due to the small sample of hospitals in the KPNC database, and the lack of patient sociodemographic diversity within the integrated network of KPNC hospitals. Confirming the validity and reliability of the measure requires data from a larger, more diverse set of hospitals and more than one EHR system. Currently there is no large national dataset that includes patient-level EHR data and captures admissions and readmissions to all hospitals from Medicare or non-Medicare claims data.

2.7 COMPARISON OF HYBRID eHWR AND ORIGINAL HWR MEASURE RESULTS

We compared the results of the original HWR measure with the results of the reengineered Hybrid eHWR to describe differences in hospital performance assessed by the two measures. We calculated the correlations between the specialty cohort specific standardized risk ratios (SRRs) and the composite risk standardized readmission rates (RSRRs) from the two models. We also compared hospitals' ranking based on the composite RSRRs calculated using the two measures. These results should also be considered preliminary given the small number of hospitals used in these analyses.

*Please see the Comprehensive Methodology Report for references and acronyms, as this is an excerpt from that report.