

TECHNICAL APPENDIX

Details of the process used to create the composite measure. We constructed risk-adjusted mortality rates for each hospital using standard methods. For patients undergoing elective AVR, let y_{ij} be a dichotomous indicator that is equal to 1 if patient j admitted to hospital i died following surgery. The risk-adjusted mortality (Y_i) for hospital i is the ratio of observed (O_i) to expected (E_i) mortality, so that:

$$(1) \quad Y_i = O_i/E_i, \quad O_i = \frac{1}{n_i} \sum_j y_{ij} \quad \text{and} \quad E_i = \frac{1}{n_i} \sum_j \hat{p}_{ij}$$

where n_i is the number of patients at hospital i , and $\hat{p}_{ij} = \text{pr}(y_{ij}=1|X_{ij})$ is the predicted probability that mortality occurred for each patient conditional on patient characteristics X . We derived the predicted probability that the outcome occurs for each patient (\hat{p}_{ij}) from a logistic regression model estimated on all patients undergoing elective AVR. The dependent variable in the logistic model is the patient's outcome (y_{ij}) and the independent variables (X_{ij}) are patient covariates potentially associated with mortality. These included patient age, gender, race, admission acuity, and comorbidities as described in the manuscript.

Our hospital-level analysis was based on a hierarchical model, in which data at the first (patient) level provided noisy estimates of hospital-level parameters at the second (hospital) level. At the first level, the mean and variance of the estimates conditional on the hospital-level parameters is:

$$(1) \quad E(Y_i | \mu_i) = \mu_i, \quad \text{and} \quad \text{Var}(Y_i | \mu_i) = V_i,$$

where Y_i is a $1 \times K$ vector of risk-adjusted mortality rates for hospital i , μ_i is the corresponding vector of underlying hospital-level quality parameters that represent the

average mortality or complication rate that a typical patient could expect at this hospital, and V_i is the $K \times K$ sampling variance-covariance matrix for the estimates in Y_i .

Note that the hierarchical nature of the data allow us to estimate V_i in a straightforward manner for each hospital, since this is simply the sampling variance of a vector of estimates derived from a sample of patients at hospital i . Our estimates of the sampling variance and covariance for these measures were derived using methods that are standard in linear models, but that are only approximations when the outcome data are Bernoulli. This is an approximation that simplifies the analysis and allows for the method to be applied to any outcome measure (whether Bernoulli or continuous). For simple composites such as Model 3 in Table 3 (combining volume and mortality for one surgery) we have found that the linear approximation yields very similar results to assuming that the data are Bernoulli.

More specifically, let $r_{ij}^k = y_{ij}^k - \hat{p}_{ij}^k$ be the patient-level residual for measure k .

We can write the k^{th} element of Y_i as:

$$(1a) \quad Y_i^k = 1 + \left(\frac{1}{E_i^k} \right) \left(\frac{1}{n_i} \right) \sum_j r_{ij}^k$$

Thus, each element of Y_i is a sample average of r_{ij}^k multiplied by a constant. We simplify by assuming that the r_{ij}^k have constant variance (σ_k^2) and covariance ($\sigma_{kk'}$) – ignoring that these depend on \hat{p}_{ij}^k for a Bernoulli – and a hospital-specific mean (the risk-adjusted mortality rate). Under these assumptions the sampling variance of Y_i^k (the diagonal

element of V_i) is $\left(\frac{1}{E_i^k} \right)^2 \frac{\sigma_k^2}{n_i}$. For two elements of Y_i from different procedures there is

no sampling covariance (the off-diagonal element of V_i is zero) because there is no overlap in patients. When two measures are estimated from the same patient sample (e.g., mortality and morbidity rates for the same procedure), the sampling covariance of Y_i^k and

$Y_i^{k'}$ is $\left(\frac{1}{E_i^k} \right) \left(\frac{1}{E_i^{k'}} \right) \frac{\sigma_{kk'}}{n_i}$. Finally, we estimate σ_k^2 and $\sigma_{kk'}$ with sample variances and

covariances in r_{ij}^k , after removing hospital-specific means from r_{ij}^k , and adjusting for the loss of degrees of freedom.

At the second level, the mean and variance of the hospital-level quality parameters conditional on observed hospital characteristics is:

$$(2) E(\mu_i) = Z_i\beta \text{ and } \text{Var}(\mu_i) = \Sigma,$$

where Z_i is a $1 \times J$ vector of observable characteristics of hospital i that are thought to be related to patient outcomes (e.g. including a constant, volume and the other structure measures), β is a $J \times K$ matrix of coefficients capturing the effect of hospital characteristic j on patient outcome k , and Σ is the variance-covariance matrix in the hospital-level quality parameters summarizing the relationships across different dimensions of hospital quality.

Estimation in our hospital-level analysis proceeded in two stages. First, we construct estimates of the variance-covariance matrix of the hospital-level quality parameters (Σ), and use this to evaluate the strength of the correlation of outcomes across the procedures. Then, using subsets of procedures for which outcomes are estimated to be strongly related, we combine information across all quality measures for these procedures to construct estimates of the underlying hospital-level quality parameters (μ_i) for each hospital. These estimates are derived from the data (Y_i, Z_i, V_i) observed for a sample of N hospitals, where in our application N is large (all hospitals in the United States performing elective aortic valve replacement on Medicare patients, approximately 1,000 hospitals).

To estimate the variance-covariance matrix of the hospital-level quality parameters (Σ), we calculated the covariance matrix of the risk-adjusted rates (Y_i), and

adjusted for sampling variability by subtracting the mean sampling-error covariance matrix (V_i):

$$(3) \hat{\Sigma} = \frac{1}{N} \sum_{i=1}^N \left\{ (Y_i - Z_i \hat{\beta})' (Y_i - Z_i \hat{\beta}) - V_i \right\} = \text{Var}(Y_i) - \text{Mean}(V_i)$$

where the coefficients ($\hat{\beta}$) on the observable hospital characteristics are estimated using a weighted least squares regression of Y on Z (weighting by n_i , the number of patients in hospital i). We estimate the i - j element of Σ weighting by the product of n_i and n_j , since these weights lead to more efficient estimates in theory and improved the accuracy of our forecasts in practice. One problem with using equation (3) is that it can obtain estimates of the variance-covariance matrix that are not positive-semi-definite (e.g., that imply correlations greater than 1). When this occurred we replaced the estimated correlation matrix with the nearest positive-semi-definite correlation matrix.

To estimate the underlying $1 \times K$ vector of hospital-level quality parameters (μ_i) for each hospital, we again used an empirical Bayes approach. The empirical Bayes estimates are a weighted average of the noisy data (Y_i) and the regression predictions ($Z_i \hat{\beta}$), where the weights depend on both the signal and noise variance ($\hat{\Sigma}$ and V_i). The equation is:

$$(4) \hat{\mu}_i = Y_i W_i + (Z_i \hat{\beta})(I - W_i),$$

where I is a $K \times K$ identity matrix, and W_i is a $K \times K$ weighting matrix estimated by

$$(5) W_i = (\hat{\Sigma} + V_i)^{-1} \hat{\Sigma}.$$

This weight is the matrix equivalent of the ratio of signal variance to total variance. Thus, equation (4) is the matrix version of a standard empirical Bayes (or shrinkage) estimator that places more weight on a hospital's own outcome rate (Y_i) when the signal ratio is high, but shrinks back toward a (conditional) mean when the signal

ratio is low. But where the usual shrinkage estimator is a weighted average of a single outcome measure and its mean, the version in equation (4) is a generalized shrinkage estimator that is a weighted average of all outcome measures and their means. Note that equation (4) yields a generalized empirical Bayes estimator for the mortality measure of each procedure that is a linear combination of the mortality and complications measures for all procedures in Y_i , along with all of the observable hospital characteristics in Z_i (such as volume) that are thought to be related to patient outcomes.

This composite measure of surgical mortality has a number of attractive properties. First, it incorporates information in a systematic way from many quality measures into the predictions of any one outcome. Moreover, if all of the estimated parameters in equation (4) were known (β , Σ , and V_i), the composite measure represents the optimal linear predictor, based on a mean squared error criterion. Since these parameters are consistently estimated as the number of hospitals increases, these composite estimates are asymptotically (in the number of hospitals) the optimal linear predictor. Finally, these estimates maintain many of the attractive aspects of existing Bayesian approaches, while dramatically simplifying the complexity of the estimation. For example, beginning with the raw patient-level data, it takes approximately seven minutes to estimate all the parameters and the composite estimates for Model 4 running STATA on a standard PC. The programs used to construct all of the estimates reported in this paper are available upon request.