Thirty-Day All-Cause Unplanned Readmission Following Psychiatric Hospitalization in an Inpatient Psychiatric Facility (IPF Readmission)

Reliability Attachment

November 2025

Submitted by:

Mathematica
P.O. Box 2393
Princeton, NJ 08543-2393
Phone: (609) 799-3535
Fax: (609) 799-0005

Contents

Reliability Attachment

Full descriptions of reliability testing methods, including formulas

Bootstrap reliability:

We computed intra-class correlation (a measure of reliability) for each hospital using a PROC NLMIXED procedure in SAS. The intra-class correlation was calculated using the following formula:

$$ICC_h = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_b^2 + \hat{\sigma}_e^2} \ ,$$

where $\hat{\sigma}_b^2$ is the between-hospital variance component which reflects the variability in the average measurements across different hospitals, $\hat{\sigma}_e^2$ is the error variance component which reflects the variability between samples *within* each hospital. Similar to signal-to-noise, the ICC quantifies the amount of variation in measure scores due to between-hospital differences rather than within.

We obtained estimates of the between-hospital and error variance components from a simple, intercept-only random effects model with no predictors. This random-effects model was fit to data made multilevel by creating 1,000 bootstrap samples, with replacement, of discharges among the hospitals meeting the minimum case threshold of at least 25 eligible discharges. By doing this, 1,000 risk-standardized readmission rates (RSRR's) per hospital were calculated. The random effects model can be expressed using the following equation and distributional assumptions:

$$Y_{ht} = \mu + \alpha_h + \varepsilon_{ht}$$

$$\alpha_h \sim N\left(0, \sigma_b^2\right); \varepsilon_{ht} \sim N\left(0, \frac{\sigma_e^2}{n_{ht}}\right)$$

where Y_{ht} is the RSRR for hospital h for bootstrap sample t, μ is a fixed effect or intercept indicating the mean RSRR, α_h is a hospital-level random effect for hospital h, and ε_{ht} is the residual bootstrap sample-level random effect for hospital h's bootstrap sample t.

With maximum likelihood estimates σ_b^2 and σ_e^2 obtained from the random effects model, we plugged these values into the formula above to compute an ICC per hospital. ICCs range from 0 to 1, where 0 indicates no agreement or reliability, and 1 indicates perfect agreement or reliability.

A high ICC indicates that the bootstrap samples provide similar measurements for each hospital, suggesting that the hospital-level metric is stable and reliable.

Table 2 provides the results of the bootstrap analysis. The mean ICC reliability was 0.66, ranging from 0.263 to 0.948.

Table 2. Accountable entity-level reliability testing results*

	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Reliability	0.66	0.26	0.35	0.48	0.56	0.62	0.66	0.70	0.75	0.79	0.83	0.88	0.95
Mean performance score	19.3%	18.1%	18.8%	19.1%	19.0%	19.2%	19.3%	19.1%	19.2%	19.7%	19.8%	19.7%	17.7%
N of entities	1410	4	140	141	141	143	140	142	140	141	141	141	1
N of persons/episodes/ encounters	278,203	100	5,471	9,093	12,810	16,448	19,462	23,547	28,836	36,327	47,590	78,619	1,277

* Reliability testing was conducted using regression adjusted scores.