NATIONAL QUALITY FORUM

Validity

NQF's Current Definition of Validity and Related Concepts

Validity refers to the correctness of measurement. **Validity of data elements** refers to the correctness of the data elements as compared to an authoritative source. **Validity of the measure score** refers to the correctness of conclusions about quality that can be made based on the measure scores (i.e., a higher score on a quality measure reflects higher quality).

Validity testing – Empirical analysis of the measure as specified that demonstrates that data are correct and/or conclusions about quality of care based on the computed measure score are correct. Validity testing focuses on systematic errors and bias. It involves testing agreement between the data elements obtained when implementing the measure as specified and data from another source of known accuracy. Validity of computed measure scores involves testing hypotheses of relationships between the computed measure scores as specified and other known measures of quality or conceptually related aspects of quality. A variety of approaches can provide some evidence for validity. The specific terms and definitions used for validity may vary by discipline, including face, content, construct, criterion, concurrent, predictive, convergent, or discriminant validity. Therefore, the proposed conceptual relationship and test should be described. The hypotheses and statistical analyses often are based on various correlations between measures or differences between groups known to vary in quality.

Validity, threats – In addition to unreliability, some aspects of measure specifications and data can affect the validity of conclusions about quality. Potential threats include patients excluded from measurement; differences in patient mix for outcome and resource use measures; measure scores generated with multiple data sources/methods; and systematic missing or "incorrect" data (unintentional or intentional).

Brief Summary of Panel Discussion to Date

- Relatively little disagreement with current definitions, but desire for more information on the "quality construct" (that is, on what the measure is trying to assess)
- Need for more insight on how higher (or perhaps lower) measure results reflect higher quality of care
- Need for additional discussion about the relationship between reliability and validity (i.e., is one needed for the other and if so, which should be considered first)
- Uncertainly about how to consider face validity when empirical validation has been conducted
- Recognition of need for deeper discussions around attribution and risk adjustment