

MEASURES MANAGEMENT SYSTEM

Risk Adjustment

The words “risk-adjusted” or “risk-standardized” often appear in the title of outcome measures used in CMS programs, as in 30-day risk-standardized mortality measures for acute myocardial infarction. Besides mortality, other measure outcomes that are commonly risk-adjusted include readmissions or complications. Process measures are not risk-adjusted; rather the target population of a process measure is defined to include all patients for whom the process measure is appropriate.

The stated purpose of risk-adjustment is to enable the accurate comparison of clinician or facility performance. Outcome measures adjust for patient characteristics that exist prior to the episode of care that may make the outcome more likely. Example characteristics might include the patient’s age, past medical history, and other diseases or conditions (comorbidities) the patient had prior to the episode of care that are known to increase the patient’s chance of the outcome.

How does risk-adjustment work?

A common formula for calculating a risk-adjusted rate for a clinician or facility is the following:

$$\text{Risk-adjusted rate} = (\text{observed rate} / \text{expected rate}) * \text{reference population rate}$$

The *observed rate* is the number of patients with the measure focus (e.g. death) divided by the number of patients in the target population (e.g. admitted with acute myocardial infarction). The target population reflects the patient characteristics of the measured entity – the clinician or facility – and the observed rate is the performance of the measure entity on that target population.

The *expected rate* is also the number of patients with the measure focus divided by the number of patients in the target population. The target population reflects the same patient characteristics of the measured entity, but the expected rate is the performance of all the measured entities in the reference population (e.g. Medicare fee-for-service) on that target population.

In the *reference population rate*, the target population reflects the patient characteristics of the reference population, and the performance of all the measured entities in the reference population on that population.

Another way to look at the formula for calculating the risk-adjusted rate for a clinician or facility is to re-arrange the terms:

$$(\text{risk-adjusted rate} / \text{reference population rate}) = (\text{observed rate} / \text{expected rate})$$

This quantity is often referred to as the “risk standardized ratio.” Assume that for a given clinician or facility the observed-to-expected ratio is 1.2. The interpretation is that the

MEASURES MANAGEMENT SYSTEM

provider's performance was 20% higher than expected for the provider's patient population. The risk-adjusted rate then assumes that the provider would have performed at the same relative performance (20% higher than expected) had the provider had the same patient population as the reference population. Note that the risk standardized ratio for the reference population is 1.0 because the observed rate is equal to the expected rate, or the risk-adjusted rate is equal to the reference population rate.

How do we know whether the assumption of same relative performance is valid?

The answer will depend on the outcome, the population, and the set of patient characteristics available and used in the calculation. Measure developers have two metrics that help decide: discrimination and calibration. Discrimination shows developers how well patient characteristics distinguish between those that had the outcome of interest and those that did not. Calibration tells how close the observed rate is to the expected rate across low risk and high-risk groups in the target population.

There are circumstances where evidence suggests a risk-adjustment model may be misleading. For example, in the case of "treatment heterogeneity" the facility or clinician quality construct is different for subpopulations of patients (e.g. mortality in younger and older patient with acute myocardial infarction). In such cases an alternative approach known as stratification (where patients are divided into two or more groups according to their expected risk of the process or outcome of interest) may be more appropriate. Careful attention to the model attributes helps to identify such circumstances, or other conditions where additional development would improve the scientific acceptability of the risk-adjustment model.

Sample Definition for a Risk Adjusted Measure.

Sometimes risk-adjustment models are estimated on samples. For example, often there is an estimation sample (where the model coefficients are calculated) and a validation sample (where those estimates are applied, and model discrimination and calibration are calculated). The sample(s) should be clearly and explicitly defined. All inclusion and exclusion criteria used to select the sample should be defined. Risk adjustment models generalize well to the extent that the samples used to develop, calibrate, and validate them appropriately represent the parent population such that the distributions of characteristics and their interactions should reflect those in the overall population.

Appropriate Time Frames.

The criteria used to formulate decisions regarding the selection of the time frame for the measure (e.g. 7-day or 30-day) should be clearly stated and explained in the measure documentation. Criteria used to identify risk factors for the stated outcomes should be clinically appropriate and clearly stated. Risk factors should be present at the start of care (e.g. within 24 hours of admission) to avoid mistakenly adjusting for factors arising due to deficiencies in care being

MEASURES MANAGEMENT SYSTEM

measured. Outcomes should occur soon enough after care to establish that they are the result of that care.

High Data Quality.

The data used for risk adjustment must be of high quality; considerations in determining the quality of data include the following:

- The data are collected in a reliable way. That is, the method of collection must be reproducible with very little variation between one collection and another if the same population is the source.
- Data must be sufficiently valid for their purpose. Validation ultimately rests on the strength of the logical connection between the construct of interest and the results of operationalizing their measurement, recording, storage, and retrieval.
- Data must be sufficiently comprehensive to limit the number of proxy measures required for the model. Obtaining the actual information is sometimes impossible, so some proxy measures might be inevitable for certain projects.
- The data collected are as recent as possible. If the measure developer were using 2012 data in a model designed to be used tomorrow, many people would argue that the healthcare system has changed so much since 2012 that the model may not be relevant.
- The data collected are as complete as possible. The data should contain as few missing values as possible. Missing values are difficult to interpret and lower the validity of the model.
- Documentation of the data sources including when the data were collected, if and how the data were cleaned and manipulated, and the data's assumed quality should be fully disclosed.

Appropriate Variable Selection.

The risk adjustment model variables should be clinically meaningful or related to variables that are clinically meaningful. When developing a risk-adjusted model, the clinical relevance of included variables should be apparent to subject matter experts (SMEs). When the variables are clearly clinically relevant, two purposes are served: the clinical relevance contributes to the face validity of the model, and the likelihood that the model will explain variation identified by healthcare professionals and/or the literature as being important to the outcome is increased. Parsimonious models and their outcome are likely to have the highest face validity and be optimal for use in a model. The strengths of the associations required to retain adjustment factors ultimately depend on the conceptual model, but are rarely a factor included in a model that is not substantively associated with the outcome variable.

Appropriate Analytic Approach.

Many factors determine an appropriate statistical model. Logistic regression or hierarchical logistic regression is often used when the outcome is dichotomous; but, in certain instances, the same data may be used to develop a linear regression model when key statistical assumptions are

MEASURES MANAGEMENT SYSTEM

not violated. Selecting the correct statistical model is absolutely imperative, because an incorrect model can lead to entirely erroneous results. The analytic approach should also take into account any multilevel and/or clustered organization of data, which is typically present when assessing institutions such as hospitals from widespread geographic areas.