

The Dirty Little Secret About Performance Measurement Data

written by Theresa Hush | February 1, 2017



“The data just hooks up.” That was an opening remark from a competitor applauding his company’s scoring system for physician quality. He went on to explain how this data produced reliable scores on quality.

The idea that data hooks up and produces a true scoring system for quality is a fantasy. Not only is data itself flawed, but it doesn’t always tell the exact truth. Treating data casually amounts to an off-hand dismissal of the complexity and inherent biases of performance measurement.

But here’s the kicker: we need to measure performance, anyway. In fact, it’s more critical than ever to measure clinical quality and costs.

Why? Because we need a starting point. [Both providers and patients require data](#) for decision-making, even if it is flawed. And, there isn’t any other way to calculate where we are on using various evidence-based protocols, what our potential cost or quality issues are, or variance

among providers. We just need to be honest about what the data really says.

Translation of performance measurement directly into scores is inappropriate. The stewards of large patient databases must safeguard the rules of performance measurement and not oversell the technology. Resulting outcomes begin the inquiry; they should not be used to turn results into immediate provider penalties.

The Truth About Data

How can data lie? To answer this question, we need to go back to how and where information is being directly captured and entered into databases.

Most identified patient data comes from one or more of three sources:

- Billing records from providers or billing companies;

- Clinical data in an EMR;

- Claims data from health plans or other payers, including all services for a population of patients for which employers, ACOs or providers of patients are responsible.

Billing data emanates from the provider and contains only those providers' services. Claims data from payers, on the other hand, will include all providers' charges for the covered patients, so that you can construct a comprehensive view of services received by a particular patient. Through the use of both provider-sourced data and claims data, more comprehensive profiles of both providers and patients can be developed.

Common among all these data sources is that each data element is entered by a human at or after a patient encounter. It then may be modified further with additional information, such as lab values obtained later, or by other individuals managing the patient or the record. The human input of the data creates a possibility of error, and the most common errors are:

- Inconsistent patient demographic data between records, including different spellings of names, wrong birth dates and gender;

- Incorrect coding of diagnoses and procedures;

- Inconsistency in the use of the EMR template itself, so that services or data are recorded differently than they actually occurred for the sole purpose of attaching a procedure code.

- Readings that vary by location of care, position of the patient, or who is performing the service. **Blood pressure is a notorious example**; wide swings in blood pressure values result from where and how the blood pressure is taken.

- Data that is missing** altogether for patients.

EMR technology may also introduce data problems. For example, codes may be out of date (yes, this frequently happens), or the EMR itself may be structured to use reporting codes (such as G-codes or Category II codes) as opposed to discrete values, so the actual “measure” values may be imprecise. Contrary to what proponents contend, getting data from an EMR does not make that data “better,” for all the reasons above. EMRs simply make data more accessible.

Finally, the transmission of data could be flawed. For a variety of reasons, the data expected to be in a particular table of the database is elsewhere. The data harvested would be therefore incorrect or missing.

Diverse Patients Also Bias Population-based Results

Aside from true data errors, comparing performance measures across providers introduces other issues. Patients have different genetic makeups as well as different risk factors and variations of disease. They come from different environments and have a varying ability to pay or manage treatment regimens. And [they may have different belief systems](#) that affect measurement results.

Grouping patients for performance will always be inexact, and no amount of risk adjustment will fix that. Again, it’s good to start an inquiry with population-based results, but applying those results to scores is not a good fit.

The evaluation of outcomes, as opposed to process measures, is, ultimately, a patient-by-patient process. Measuring population-based results provides questions but not enough answers to serve as “scores” of quality. This means that there must be a capacity within performance measurement and improvement for input by the provider and/or patient so that the results can be explained.

Measures Have Problems, Too

If all these issues weren’t enough to dispel an idea of the perfect measurement system, take a look at the measures themselves. Most performance measures are still “process” and not outcome measures. What they are measuring is whether or not a medical service was provided. For example, there is a measure of how many patients meeting eligibility criteria have colon cancer screening. Patients may get this service anywhere, so the measurement is always insufficient to explain what the provider’s results are across a population. This and similar measures for preventive services are almost always incorrect if taken literally.

In addition, outcome measures themselves can be problematic. For example, recent research

into one of the most common outcome measures for patients with diabetes, HgBA1c levels, finds that the [actual levels may be less important than the medication used to control hemoglobin levels](#).

Also important is the absence of measures related to outcomes—complications, redos and infections are just a few surgical measures left out of all but customized programs. Physicians largely do not consistently code such problems, and yet they are clearly indicators of trouble.

Some measures also create unexpected results, such as inappropriately higher-cost follow-up services. A few years ago, ductal carcinoma in situ was demoted as a form of breast cancer that always required treatment. Improved imaging technology, coupled with higher rates of measurement, had an [unintended effect of increasing the number of cancer diagnoses and surgeries](#).

The Measurement Process Itself Compounds Errors

The truth of the matter is that the following are inherent weaknesses in the process:

Providers want something quick and dirty, mainly because they don't see measures as meaningful. They often don't have the time or interest to even look at their data.

Payers are looking for a label of quality to demonstrate concern, and the accuracy of the information is secondary.

Hidden financial interests corrupt the measures themselves. When the research backing up guidelines/measures may not be fully reported, how can we blame providers for lack of commitment?

Who pays for provider measurements? Providers. This is backward; the measurement should be more independent.

Provider fear of measurement fosters extreme reactions that may undercut the measurement process, depending upon the penalties, rewards or public disclosure involved.

How to Make Performance Measurement More Accurate *and* More Relevant

Regardless of limitations, [measuring performance is now a necessary process](#) in the industry. It usually has widespread support from health systems (although not necessarily providers), and is required by Medicare and many private payers.

But how do we make these efforts more meaningful and relevant?

Focus on measures of clinical significance to providers. Measures are often delegated to administrators and other office staff because physicians considered this a “Mickey Mouse” process task. To engage providers in real progress, focus on outcome measures that matter.

Lay out the measures in a performance improvement inquiry, rather than using them as scores. The fact is that we don’t know what accounts for variation between providers or patients, and thus we should be questioning that legitimacy.

Collaborate instead of legislate. If performance measurement is meant to lead to improvement, it should not be part of a scheme to legislate behavior. Care processes and outcome measures should be distinct, and outcome measures—which can include cost-related outcomes—deserve the investigation of alternatives for improvement.

Seek more data. The questions raised about the performance must be answered, and it will almost certainly lead to more data, probably from patients or their caregivers. That’s normal and it is an opportunity to involve those patients in improving their results.

Measure provider engagement, collaboration and responsiveness. Make it a top priority to help providers focus on their data and participate in improvement projects.

Share results among teams. Improving performance requires the sharing of success stories and tactics that work. [Change is social](#), and the technology you are using should be able to accommodate a dynamic process.

Founded as ICLOPS in 2002, Roji Health Intelligence guides health care systems, providers and patients on the path to better health through [Solutions](#) that help providers improve their value and succeed in Risk. Roji Health Intelligence is a CMS Qualified Clinical Data Registry.

Image Credit: [Ryan McGuire](#)