QSEA Quality Measurement & Pitfalls in Data Interpretation

Form groups of 2 or 3 at your tables. Read the scenario and discuss the questions which follow. We will ask small group members to report on their discussion during a large group debriefing.

You have been asked to engage internal medicine residents in a QI project. The residents want to work on improving care for diabetic patients in their clinic. A diabetic care improvement team is assembled and includes 5 residents, 2 clinic nurses, 2 clinic certified nurse assistants (CNAs – they room the patients in this practice), an information systems manager, and an attending general internist. You wish to have the residents take the lead on the project as much as possible and their schedule allows them to do so.

1. What process, outcome, and balancing measures are needed to evaluate performance? For this exercise, you may wish to limit the number of measures you choose to 3. Assuming there is room for improvement, what measures might help you identify drivers of suboptimal performance?

   Residents typically underestimate the importance of critically thinking about measures early in the course of a QI project. Having them work through the details of a set of measures helps them appreciate the complexity of quality measurement, the pitfalls related to poorly designed accountability measures, and critically appraise QI publications and presentations.

   In this case, it's unlikely that there will be a structure measure. Potential structure measures might include a dedicated dietician or diabetes educators in the clinic. Process measures might include % of patients who have a HbA1C, lipid profile, or microalbumin checked, % patients referred to ophthalmology, % of patients who actually have a dilated eye exam, % who have a foot exam. Outcome measure might be % with HbA1C≤9%, LDL<100, last SBP<140 and DBP<90. Teachers should point out that these are intermediate outcome measures. The long term outcomes related to these measures are incidence of nephropathy, neuropathy, retinopathy, myocardial infarction, and stroke.

2. How will you acquire this data? Does it need to be manually abstracted from charts, generated as a report from an information system, etc? Which patients are included in your measures? Are there exclusion criteria?

   Residents may have a number of reactions when making efforts to obtain QI data. They may be amazed at how much data already exists and may know nothing about ICD-9 codes, present on admission indicators, etc. On the other hand, they may become frustrated not knowing how to acquire existing data, not having established connections to data analysts, not having advanced analytical skills, or realizing that the data they really really want will require work (like manual chart abstraction).

   These measures need to include specifications for time period, inclusion, exclusion, and attribution. For example, % patient with lipid profile checked who had 2 or more visits in the clinic over the past year, excluding patients >75 . Another potential exclusion would

be stage 4 cancer or enrolment in hospice, but this information may be harder to come by. This outpatient diabetes example is nice in that NQF endorsed measures already exist for diabetes and many providers and/or groups are collecting this data as part of the Center for Medicare and Medicaid Services (CMS) Physician Quality Reporting (PQRS) incentive program. Therefore, technical specifications already exist for diabetes measures but the residents may wish to revise the specifications and/or add additional measures. Some groups use manual chart abstraction for these measures, some use special codes assigned at the point of care, some leverage their electronic health record in get this data. If using the EMR, it may be difficult to automate the collection of some data (foot exam is the best example).

Teachers should help residents understand some of the potential unintended consequences of using HbA1C as a threshold (or achievement) measures, especially in the context of incentives (pay for performance). Imagine two patients: one starts out with the HbA1C of 9.1 and later achieves a HbA1c of 8.9 while the other starts out with a HbA1c of 12.9 and later achieves a HbA1c of 9.1. We helped the second patient much more, but we don't get credit. If strong financial incentives are in place, a provider or group may be inclined not to accept (or schedule for follow up) patients who seem especially complex and/or non-adherent. Creating a measure which credits either achievement or improvement would be better.

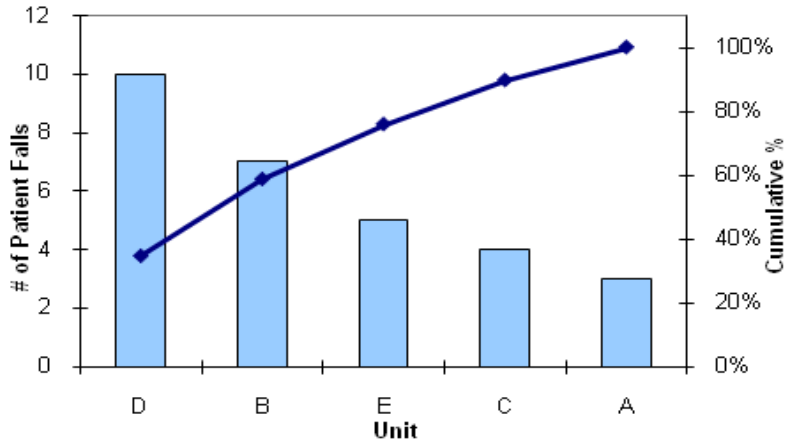
Drivers for poor performance for some diabetes measures might be accomplished through chart review. For example, a patient may not have an LDL checked in the resident clinic because they see a cardiologist at an outside institution that is checking their LDL.

3. How will you analyze and interpret the data and performance over time? Will you use certain types of charts?

   Residents may underestimate the difficulty in analyzing QI data as well and typically have no prior experience distilling complex data into a simple story for stakeholders. Residents will need to know who they can turn to for help. Depending on the program, a resident may get help from someone in the quality department or from someone in a research group.

   A pareto chart could help explore the more common reasons for suboptimal performance. A pareto chart helps visually display the pareto principle: 80% of the variation in performance can be explained by 20% of potential causes/factors. A pareto chart plots frequency of events (y axis) in descending order by category (x axis) and also include a cumulative % of events. One has to be careful in establishing categories and

using a pareto chart – ideally categories should be mutually exclusive. Example below:



An outpatient project like this is challenged by the fact that most of our measures should occur only once or a handful of times per year. It would be important to look several years back to see whether improvement was occurring even before we start our project. Some process measures would be feasibly examined in shorter time frames on a rolling basis. For example, the team could analyze the % diabetic patients seen each month with documented foot exam during the visit.

Outcomes again, like HbA1c, LDL, and blood pressure may take longer to improve. Process measures could be created which should translate into improvement. For example, % of diabetic patients with LDL>100 who had a change in the cholesterol lowering medication made (either a new Rx, higher dose, or change to more potent medication). Run or control charts would work nicely to track these measures over time.

QSEA Quality Measurement & Pitfalls in Data Interpretation

Form groups of 2 or 3 at your tables. Read the scenario and discuss the questions which follow. We will ask small group members to report on their discussion during a large group debriefing.

You have been asked to engage internal medicine residents in a QI project. The residents want to work on reducing nosocomial venous thromboembolism (VTE). A VTE reduction team is assembled and includes 3 residents, 2 staff nurses, a nurse manager, a hematology fellow, a pharmacist, an information systems manager, and a hospitalist. You wish to have the residents take the lead on the project as much as possible and their schedule allows them to do so.

1. How would you guide residents (and other project team members) into determining what process, outcome, and balancing measures are needed to evaluate performance? Similarly, how would you guide residents in determining which measures might explain drivers of suboptimal performance?

   Residents typically underestimate the importance of critically thinking about measures early in the course of a QI project. Having them work through the details of a set of measures helps them appreciate the complexity of quality measurement, the pitfalls related to poorly designed accountability measures, and critically appraise QI publications and presentations.

   In this case, it's unlikely that there will be a structure measure. An order set or clinical decision support prompt could be considered a structure measure and would also be a potential intervention. Process measures might include DVT risk assessment completion and orders consistent with recommendations based on risk assessment. Process measures may also include rates of pharmacologic prophylaxis (which may or may not be stratified by appropriateness). A balancing process measure might be rates of pharmacologic prophylaxis orders for low risk patients and/or patients with contraindications. Outcomes include DVT/PE rates and a balancing outcome measure may include hemorrhage. Drivers of poor performance might include analyses based on hospital unit, service, person who completed the risk assessment, cessation of pharmacologic prophylaxis for invasive procedures (and perhaps not restarted), etc.

2. How will residents acquire this data? Does it need to be manually abstracted from charts, generated as a report from an information system, etc? Who will obtain the data? Which patients are included in the measures? Are there exclusion criteria?

   Residents may have a number of reactions when making efforts to obtain QI data. They may be amazed at how much data already exists and may know nothing about ICD-9 codes, present on admission indicators, etc. On the other hand, they may become frustrated not knowing how to acquire existing data, not having established connections to data analysts, not having any analytical skills, or realizing that the data they really really want will require work (like manual chart abstraction).

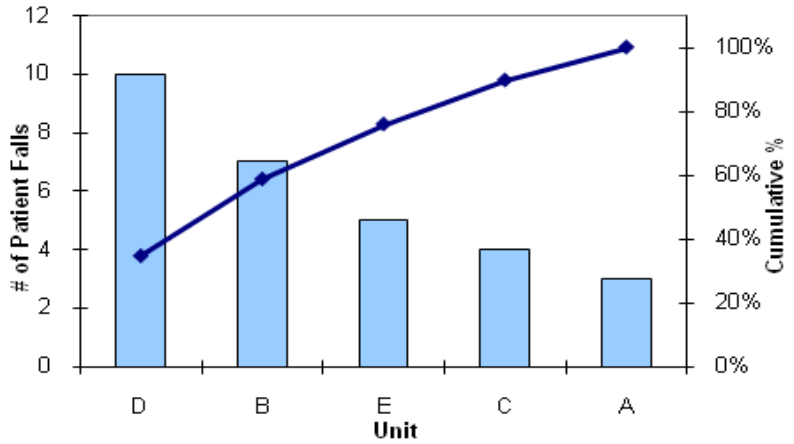
   Risk assessment could be built into admission order sets (paper or electronic). A report might be created for risk assessment completion and orders consistent with

assessments. Pharmacologic prophylaxis rates could be stratified as total (# on Rx proph/all pts) but even better would be having it broken out as only those with indications and without contradindications (# on Rx proph/# pts with indications and without contradindications). Patients on full dose anticoagulation for another reason could be excluded from numerator and denominator. Outcomes may be difficult. The team could use billing data to identify patients with DVT/PE ICD-9 codes in a secondary diagnosis field and not designated as present on admission. AHRQ PSI has an algorhithm and list of relevant ICD-9 codes for post-op VTE which could be adapted as well. The PSI algorhithm excludes patients with a VTE code in the primary diagnosis field. Keep in mind that coded data should be confirmed for accuracy using a chart review on a sample of patients. Also, outcome data should ideally be risk adjusted. Risk adjustment likely goes beyond the scope of what most QI teams can do. If there is no reason to think that patient mix has changed, a run or control chart of VTE as an outcome is reasonable. Some external data vendors will perform risk adjustment – for example UHC will do this for postoperative VTE (PSI-12, which of course is outside the scope of this project).

3. How will you guide residents in analyzing and interpreting the data and performance over time? Will residents use certain types of charts and how might you teach them which to choose?

   Residents may underestimate the difficulty in analyzing QI data as well and typically have no prior experience distilling complex data into a simple story for stakeholders. Residents will need to know who they can turn to for help. Depending on the program, resident may get help from someone in the quality department or from someone in a research group.

   A pareto chart could help explore the more common reasons for suboptimal performance. A pareto chart helps visually display the pareto principle: 80% of the variation in performance can explained by 20% of potential causes/factors. A pareto chart plots frequency of events (y axis) in descending order by category (x axis) and also include a cumulative % of events. One has to be careful in establishing categories and using a pareto chart – ideally categories should be mutually exclusive. Example below:



Analyzing adherence to the process measure (# on Rx proph/# pts with indications and without contradindications) is complicated by the fact that patient's risk for nosocomial VTE and their contraindications for Rx proph may change throughout out the hospital stay. One way to analyze this is to restrict analysis to specific days (day #2). Another is to analyze the data by patient-days. Run or control charts would work nicely to tracking these process measures over time.

Outcomes again, may be difficult to assess but a run and or control chart, assuming no change in case mix, may work. Enough time will need to be given to see a change as the number of events (VTEs) should be low.