NATIONAL QUALITY FORUM

National Voluntary Consensus Standards for Patient Outcomes
Summary of the Surgical Advisors Conference Call
March 31, 2010

Members participating: Lee Fleisher, MD (Chair); David Herman, MD; David Penson, MD, MPH; Joe Putnam, MD; Robert Karpman, MD, MBA; Rocco Ricciardi, MD, MPH

NQF Staff participating: Reva Winkler, MD, MPH; Sarah Fanta; Hawa Camara

Measure Steward Representatives: Bruce Hall, MD (ACS)

Review of Measures

OT1-015-09 Risk Adjusted Case Mix Adjusted Elderly Surgery Outcomes Measure (American College of Surgeons)

Measure Developer's Comments:

- The elderly are a special population with a higher risk of complications that are a great burden to patients and society; they appear to be a special population with different outcomes.
- The measure includes 3000 procedures/CPT codes across a variety of surgical specialties.
- This is a dedicated risk model that relies on a limited number of data elements, developed from experience with National Surgical Quality Improvement Program (NSQIP). It is very parsimonious but still performs well; it is estimated that 175-180 eligible cases are needed over a course of a year to reach an acceptable level of reliability.
- Risk adjustment model is standardized by CPT groupings--135 CPT groups, initially run as a dummy against regression in preliminary that generates odds for CPT group. The variable is re-entered for individual institutions to standardize the risk grouping so there is no problem for valid comparison among institutions. Reliability may increase as the number increases this number is achievable by 90% of hospitals in the country. Small hospitals that do not do an adequate number of procedures (10%) will not be included. Ninety-five percent of all of the cases are performed within 90% of hospitals with sufficient case loads.

IMPORTANCE TO MEASURE AND REPORT		
1a impact	Completely	High impact, there is data to prove that the opportunity to close gaps exist. Morbidity and mortality are addressed; therefore, this is a very important measure. This measure uses a subset of NSQIP; outcomes considered to be most important, chosen based on the burden (frequency) or severity.
1b gap	Completely	
1c relation to outcomes	Completely	
SCIENTIFIC ACCEPTABILTY		
2a specs	Completely	The definitions for NSQIP are clear and concise and based on clinical data. The only question is how to determine functional status. Developer noted that this information is usually documented well in the nursing assessment using ADLs. Data element shows high degree of reliability. Reliability estimate of
2b reliability	Completely	
2c validity	Completely	
2d exclusions	Partially	
2e risk adjustment	Partially	

2f meaningful differences	Partially	0.4 is low. Denominator includes 3000 CPT codes/procedures across multiple specialties. A systematic sampling method is described to avoid selection bias. Exclusions included ASA 6, trauma, and transplant procedures - hip fracture is considered a traumatic event that would be excluded. Concern about excluding this large population of elderly patients. Suggestion to separate trauma from injuries because, particularly for the elderly patient, minimal trauma to create a fracture should either be a separate measure or included in this. By excluding fractures for this elderly population, several hundred thousand cases are being excluded from this measure. Risk model controls for surgical case mix by standardizing for CPT groups. It is unclear how accurately the CPT grouping can risk stratify for surgical case mix. Risk model - General estimation equation, not a fully hierarchical model but has similar characteristics. Good performance statistics. The model is simple with a limited number of factors. Disparities – race and ethnicity not included in the model; NSQIP does collect race and ethnicity so stratification is possible. Socioeconomic status can only be determined by zip code or geographic proxy and not according to payment.
2g comparability	N/A	
2h disparities	Completely	
USEABILITY		
3a distinctive	Partially	While using O/E ratio seems meaningful for providers, it is not known how this information will be used in public reporting. Public reporting is different than the quality improvement – the facts are there for quality improvement. For harmonization, other measures are part of NSQIP; no other measures are applicable to a broad surgical population. Additive value-sheds light on an important population; the global measure is very applicable for individuals going to a hospital.
3b harmonization	N/A	
3c Added value	Completely	
FEASIBILITY		
4a data a byproduct of care	Partial	A limited number of hospitals belong to NSQIP and the TAP had concerns regarding the implementation plan for broader use. The developer noted that these measures have been developed in conjunction with CMS – CMS is likely to implement similarly to the SCIP measures (i.e., the data would be sent to a central location for calculation and then reported back). To date, the
4b electronic	Minimal	
4c exclusions	Complete	
4d inaccuracies/errors	Partial	

4e Implementation	Partial	measure has not been used outside of NSQIP. There is a data collection burden to abstract some clinical elements and follow-up with patients for any complications within the 30 days occurring after discharge – estimated at 1/10 to 1/3 FTE. Some data elements can come from administrative data – including the exclusions – but clinical elements require abstraction. Some judgment is required for the clinical elements such as functional status. This would be an added function for hospitals that do not currently have the resources. Electronic sources are not currently available in many hospitals it is dependent on whether there is a structured medical record. The data is entered by hospital employees and unless there is an audit it is nearly impossible to identify inaccuracy or error.

Recommendation: The surgery advisors were particularly concerned with exclusion of elderly patients with hip fracture which is a common surgery in this population. The advisors strongly recommend that NSQIP revisit the exclusion of patients with minimal trauma such as hip fracture. The developer noted that they do not collect the data. Alternatively, members would like to see a similar measure for the hip fracture population.

Public comment - none