Part 2: The American Society for Metabolic and Bariatric Surgery Bariatric Surgery Quality Improvement Program (ASMBS NBSQIP) – On what outcome measure should the ASMBS base accreditation?
By Robin Blackstone, MD, President, American Society for Metabolic and Bariatric Surgery

“The ASMBS has embarked on an evaluation of our current BSCOE program. Throughout the last 10 months, many of our colleagues have been working in ASMBS committees and sub-committees to evaluate different parts of the current program and make proposals for an evolution of the program. Those proposals are now ready for member comment and input. The second installment details the scientific arguments and data for using different outcomes measures. Once all of the information has been presented we will have a survey/comment period so that you can give the leadership your feedback.”

Robin Blackstone, MD

Introduction
Throughout the history of surgical intervention, surgeons have been keenly interested in improving patient safety. The ASMBS has historically focused its quality measurement efforts on accreditation and ensuring that bariatric surgery is performed in safe/appropriate settings. This function remains paramount in importance and ASMBS is committed to using measures that best reflect hospital outcomes and true quality. In addition, ASMBS leadership recognizes the need for quality measurement to support not just accreditation, but quality improvement in all settings where bariatric surgery is delivered.

In the classic Donabedian paradigm for assessing quality of care, three measures prevail as indicators of quality: Structure, Process of Care and Direct Outcomes.\(^1\)

There are strengths and weaknesses of each measure as outlined in the following Table 1 from ACS Surgery: Principles and Practice Performance Measures in Surgical Practice.\(^2\)

<table>
<thead>
<tr>
<th>Type of Measure</th>
<th>Examples</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| Structural      | Procedure Volume  
Intensivist-managed ICU | Measure is expedient and inexpensive  
Measures are efficient - a single measure may relate to several outcomes  
For some procedures, structural measures predict subsequent performance better than process or outcomes measures do | Number of measures is limited  
Measures are generally not actionable  
Measures do not reflect individual performance and are considered unfair by providers |
Process of Care

| Appropriate use of prophylactic antibiotics |
| Measure reflect care that patients actually receive—hence, greater buy-in from providers |
| Many measures are hard to define with existing databases |

Measure reflect care that patients actually receive—hence, greater buy-in from providers

Many measures are hard to define with existing databases

Measure are directly actionable for quality—improvement activities

Extent of linkage between measures and important patient outcomes is variable

For many measures, risk adjustment is unnecessary

High-leverage, procedure-specific measures are lacking

Direct outcome

| Risk-adjusted mortalities for CABG from State or national registries |
| Face Validity |
| Sample sizes are limited |

Face Validity

Measurement may improve outcomes in and of itself (Hawthorne effect)

Clinical data collection is expensive

Concerns exist about risk adjustment with administrative data

| Table 1: Primary Strengths and Limitations of Structural, Process and Outcomes Measures |

Once accurate data is collected, it can be analyzed to provide information that can be turned into improvement in quality. The science of measuring variability and predicting the future performance of a facility or hospital has advanced. The important question is where does the variation come from? Iezzoni attributes variation to her “algebra of effectiveness” meaning variation in outcomes is attributed to one of three factors: chance, case mix and quality of care.\(^{(3)}\)

This model has evolved as the understanding of contributory processes has evolved. A more current concept is presented below\(^{(4)}\):

---

![Conceptual model of relationships between structure, process of care, complications, and mortality after surgery.](image)

**Figure 2**

Current relationship of ASMBS BSCOE Standards to the Quality Paradigm

In order for a program to become a fully approved BSCOE, each program goes through a provisional step where they begin to align their program with the 10 requirements for achieving full approval. The biggest barrier to attaining full approval BSCOE status appears to be requirement #2, the volume of cases, although specific details of other requirements may also be a barrier.
Throughout time, the Bariatric Surgical Review Committee (BSRC) has been called upon to interpret the standards for individual programs/surgeon situations; some requirements have become increasingly restrictive and prescriptive of surgeon/facility behavior. Some of the interpretation of the original standards by the BSRC is based on best medical opinion. In some cases, additional requirements have added a further cost burden to the facility and surgeon without a clear demonstration of increased value or quality. An example of this was the addition to requirement #4: Programs were required to provide angiography capabilities at their facility. For large hospital systems, this did not present an issue as they already had this capability. For smaller size or rural hospitals, it was a barrier for participation in the BSCOE program. The angiography requirement was recently reviewed by the Rural Subcommittee with a recommendation that it be removed from the requirements, as there was no evidence basis to support it. The Quality and Standards Committee and the Executive Council concurred and it was removed in July 2011.

The primary proxy for quality in the ASMBS BSCOE program is volume. Table 3 below outlines the strengths and weaknesses of the current ASMBS BSCOE model.

<table>
<thead>
<tr>
<th>Type of Measure</th>
<th>Standard</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural</td>
<td>Procedure Volume</td>
<td>Supported in multiple publications as a proxy for patient safety</td>
<td>See discussion below</td>
</tr>
<tr>
<td></td>
<td>Appropriate equipment and instruments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process of Care</td>
<td>Clinical pathways (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Qualified call coverage</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patient support groups (Grade D Evidence) (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct outcome</td>
<td>Surgeons and programs report mortality and adverse events, but these results are not currently used in evaluating whether a program qualifies for the ASMBS COE.</td>
<td>Allows programs some leeway in performance since no standard has been established.</td>
<td>Creates a situation where programs may meet criteria but have poor outcomes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Statistical noise because of low volume of cases in each individual center</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>See discussion of risk adjustment</td>
</tr>
</tbody>
</table>

Table 3: Relationship of current ASMBS quality construct to type of quality measures

Conclusion: Currently a program can have met all the standards and the interpretations of those standards, including reporting their outcomes, but not actually be performing bariatric surgery within reasonable rates for mortality and morbidity. The major opportunity for change within the current paradigm is for the program to evolve to having a predictive outcome measure of performance.
On What Measures should the ASMBS base accreditation?

Risk Adjustment
Since the inception of the ASMBS BSCOE program, the intent was to eventually use risk-adjusted outcomes to define and predict quality.

Risk adjustment using a rich clinical database with large volume of patients seems intuitively important. Surgeons want to believe that if their results are poor in comparison to established norms that it is because they are operating on sicker patients. Critical to that analysis, however, is that the data collection must include all adverse events. The farther away in time from initial hospitalization you try and collect data, the less reliable it is. This is magnified by other issues with collecting data that will be detailed in the last segment of the conversation about quality.

In order for risk adjustment to accurately predict future adverse events, complications must occur with sufficient frequency to be analyzed statistically. In addition, the complications need to occur in a predictable pattern, as in cardiac surgery. There are several older trials that have completed a correlation analysis to identify risk factors associated with mortality or other complications in bariatric surgery. The variability of these analyses speaks to the low frequency of serious complications, the lack of consistent correlation of specific risk factors with these complications and, to some extent, variable statistical analyses and approaches.

The resulting summary (Table 4) shows the impact of the procedure type and patient factors that are important in widely variable reports of risk adjustment from different eras of surgery.

<table>
<thead>
<tr>
<th>Author/Date</th>
<th>Description of Study</th>
<th>Risk Factor</th>
<th>Adjusted Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gupta et al J Am Coll Surg 2011&lt;sup&gt;(7)&lt;/sup&gt;</td>
<td>Bariatric NSQIP (All bariatric procedures including revisions) 11,023 Patients Endpoint: Selected 17 postoperative complications</td>
<td>MI/Angina  Dependent Functional Status Stroke Bleeding Disorder Hypertension  BMI 35&lt;45  BMI 45-60 Procedure Type: BPD/DS OGBP LGBP</td>
<td>3.65: CI 1.23-10.8  3.48: CI 1.78-6.80  3.01: CI 1.09-7.67  2.37: CI 1.47-3.38  1.34: CI 1.10-1.63  0.9: CI 0.67-1.21  0.69: CI 0.52-0.91</td>
</tr>
<tr>
<td>Nguyen et al Surgery 2011&lt;sup&gt;(8)&lt;/sup&gt;</td>
<td>Nationwide Inpatient Sample 304,515 Patients 2006-2008 Endpoint: In-hospital Mortality Overall 0.12%</td>
<td>Male Gender Age &gt;50 Congestive Heart Failure Peripheral Vascular Disease Chronic Renal Failure Open procedure GBP</td>
<td>1.7: CI 1.2-2.2  3.8: CI 2.8-5.0  9.5: CI 6.8-13.2  7.4: CI 4.5-12.2  2.7: CI 1.6-4.5  5.5: CI 4.4-7.2  1.6: CI 1.2-2.4</td>
</tr>
<tr>
<td>Finks JF et al MBSC Ann Surg 2011&lt;sup&gt;(9)&lt;/sup&gt;</td>
<td>MBSC 25,469 Patients All Procedures</td>
<td>Most significant risk factor was procedure type: Duodenal Switch Laparoscopic Gastric Bypass</td>
<td>9.68: CI 6.05-15.5  3.58: CI 2.79-4.64</td>
</tr>
</tbody>
</table>
June 2006-December 2010
Endpoint:
Grade 2 or 3 Complications
Open Gastric Bypass
Sleeve Gastrectomy
Patient Factors:
Previous History VTE
Mobility Limitations
Coronary Artery Disease
Age over 50
Pulmonary disease
Male gender
Smoking history
3.51: CI 2.38-5.22
2.46: CI 1.73-3.50
1.90: CI 1.41-2.54
1.61: CI 1.23-2.13
1.53: CI 1.17-2.02
1.38: CI 1.18-1.61
1.37: CI 1.15-1.64
1.26: CI 1.06-1.50
1.20: CI 1.02-1.40

Table 4 Summary of the currently published risk adjustment publications
MI = myocardial infarction; OGB = open gastric bypass; LGBP = laparoscopic gastric bypass; LAGB = laparoscopic adjustable gastric band; BPD/DS = biliopancreatic diversion/duodenal switch; BMI = body mass index; GBP = gastric bypass
*Pulmonary Embolus Risk = history of previous venous thrombosis, pulmonary embolus, inferior vena cava filter, right heart failure and obesity hypoventilation

Risk adjustment is important and should be used in the development of a quality measure despite limitations including no single patient-derived risk factor has emerged as a predictor of complications and the incidence of serious complications is a J-shaped curve. There has been no published risk adjustment data from BOLD. The impact of the procedure itself on risk is substantial. In the current quality paradigm, this very important issue is not taken into account.

Volume
There are many reports of volume as a surrogate for quality. In the early days of the ASMBS BSCOE program, reports of the importance of volume to quality were cited to justify using volume as the primary quality indicator in the program. Hence, as the standard in regard to surgeon experience, the volume requirement for surgeons was set at 50 cases and for the facility at 125 cases.

One study attempted to determine an appropriate volume requirement by an ASMBS COE surgeon. The Longitudinal Assessment of Bariatric Surgery (LABS) is a multicenter, prospective trial to study 30-day outcome data on patients undergoing bariatric surgery. Mortality is a rare event in bariatric surgery precluding the use of mortality alone as an endpoint. LABS developed a composite event (CE) endpoint that includes the occurrence of at least one of the following events: death, venous thrombosis, pulmonary embolism, reoperation, non-discharge at 30 days and repeat hospitalization within 30 days after initial discharge. Risk adjustment in LABS is based on patient BMI, functional status, history of deep vein thrombosis and history of obstructive sleep apnea.

After adjusting for patient risk, the effect of surgeon volume on outcomes for RYGP procedures in LABS showed that for each increase by 10-cases per year in surgeon volume, the rate of composite events improved by 10 percent. No significant differences were observed in mortality between low and high volume surgeons. Unfortunately, the study was not powered sufficiently to detect small differences. An additional contribution of this study is the demonstration that the risk of a serious complication in the hands of a low-volume surgeon is greatly exaggerated in a higher risk profile patient. The observed relationship between surgeon RYBP volume and CE rates was continuous, illustrating that there was no satisfactory level of annual case volume that could act as a threshold for surgeon credentialing within the BSCOE.
Additional reports in the field of bariatric surgery as well as other specialties have validated the volume and outcome relationship and reports on surgeon volume alone have also been published. In 2003, Courcoulas et al using an administrative database in Pennsylvania reported that surgeons performing <50 RYBP cases annually had a significantly increased rate of complications. The effect was worse if a low volume surgeon was performing in a low volume facility. In 2006, a study from New York State using an administrative database looking at RYBP and gastroplasty patients found that both surgeon volume <100 cases annually and low facility volume of less than 150 cases were associated with increased risk of complications.

Reports from administrative data focus primarily on hospital volume have also shown a relationship between procedure volume and outcomes.

The Michigan Collaborative data found an inverse relationship with volume and complication rates. The following table shows the relationships between Surgeon and facility volume.

### Annual bariatric procedures by surgeon

<table>
<thead>
<tr>
<th>Surgeon Volume</th>
<th>&lt;100</th>
<th>100-249</th>
<th>&gt;=250</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk adjusted serious complication rate</td>
<td>3.8 (3.2-4.5)</td>
<td>2.4 (2.1-2.8)</td>
<td>1.9 (1.4-2.3)</td>
</tr>
</tbody>
</table>

**Annual Bariatric Surgery Complications by Surgeon Volume in MBSC**

### Annual bariatric procedures by hospital

<table>
<thead>
<tr>
<th>Hospital Volume</th>
<th>&lt;150</th>
<th>150-299</th>
<th>&gt;=300</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk adjusted serious complication rate</td>
<td>4.1 (3.0-5.1)</td>
<td>2.7 (2.2-3.2)</td>
<td>2.3 (2.0-2.6)</td>
</tr>
</tbody>
</table>

**Annual Bariatric Surgery Complications by Hospital Volume in Michigan**

In recent years, the value of volume alone as the determinate of quality has been questioned, especially in procedures in which the mortality rate is low. Coronary artery bypass graft (CABG) is one procedure where more recent data has called volume into question as a surrogate for quality. In a report by Paul Kurlansky, MD, at the annual meeting of the American Association for Thoracic Surgery in 2011, a comparison of low-volume centers (less than 200 cases per year) with high-volume centers (200 cases or more) demonstrated no significant difference in mortality for either surgeon or facility volume. How transferrable this is to bariatric surgery is unknown, as this volume even at less than 200 cases annually would be considered high in our field of surgery. The Society for Thoracic Surgeons, always a leader in quality, has adopted a composite score for quality and adopted process measures through the National Quality Forum. Lack of compliance with the NQF measures, in the same data presented by Dr. Kurlansky, was significantly and highly predictive of morbidity regardless of volume even after adjustment for patient risk factors.

Another way of looking at volume and mortality is demonstrated in a paper by Ghaferi. Using the ACS NSQIP clinical database, he studied 84,730 patients undergoing general surgery and vascular procedures. Although mortality rates differed from high-volume to low-volume centers, 3.5 percent to 6.9 percent, major complication rates were similar in the high-mortality (24.6 percent) and low-mortality hospitals (26.9 percent). The failure of the lower volume center to “rescue” the patient from the complication appears to account for the difference in the mortality. Dr. Nguyen reported at the 2011 ACS Clinical Congress (General Surgery News December (2011) 38:12), looking at 35,000 bariatric operations performed between October 2007 and December 2009 that mortality at accredited centers was 0.06 percent compared with 0.21 percent at non-accredited centers; however, the mortality was linked to the ability of accredited centers to rescue the patient rather than associated with the volume of cases since the complication rates were similar.

This represents a major opportunity that could be leveraged to improve mortality if the ASMBS quality program is reoriented toward process improvement and sets up a network of prearranged transfers if a patient gets in trouble in a location that does not have the resources to rescue, like 24-hour critical care teams. This would also allow patients to be treated for their primary procedures within their local area and likely expand access.

Finally, there are a few central problems with volume as the sole discriminator in the current construct of the ASMBS BSCOE. Based on the only prospective study (LABS), the ASMBS may not know the level of volume (surgeon or facility) adequate to achieve good outcomes. The current volume requirements do not discriminate between procedure types and count RYBP and AGB the same for the volume requirement. Every risk adjustment report emphasizes the
importance of procedure type in the risk adjustment. If bariatric surgery is limited to facilities that have high volume, then access will be severely limited, and patients will continue to have to travel outside their medical home to have a procedure making it less likely to receive necessary follow up. If the patient experiences a complication, then they are often treated by a local general surgeon who is caring for a medically complex patient in the most difficult of situations. An additional problem with surgeon volume is that surgeons may practice at BSCOE and non-BSCOE hospitals. It is unlikely that BOLD records this with any accuracy or captures the adverse events even during the index hospitalization, despite an attempt to do so.

Clearly volume is an important variable in achieving safety, but it should only provide a portion of the quality “signal” necessary to develop the composite quality measure used to discriminate between programs.

**Recommendation:** Include volume in the accreditation standard within the composite measure to the extent indicated by the data rather than as an exclusionary mechanism. Work to include all programs/surgeons practicing bariatric surgery within the new program by allowing programs to enter with no minimum volume requirement.

**Adjustment of hospital-reported data for reliability.**
Statistical modeling predicts that outcomes reported by facilities may occur due to chance depending on the volume. For instance, a hospital with an annual volume of 1,000 cases reporting mortality of 2/1000 is probably a better representation of the true risk of death than a small hospital reporting one death in 80 patients. To reduce this statistical “noise” in the data, a technique has been written about and utilized extensively called “reliability adjustment”. Reliability is a measure of precision and is a function of hospital sample size and the amount of true variation across hospitals. The overall observed effect is shrunk back toward the mean of the facilities with similar volume (not the overall mean) thereby correcting the observed risk-adjusted rate by the volume of the facility. This allows for all hospital volumes within the sample to be assigned a reliability adjustment factor (from zero to 1.0) and using empirical Bayes techniques, the observed-to-expected ratio is adjusted to reduce the statistical “noise” in the sample size.\(^{(25)}\) For a short slide presentation that shows how this works, please go to: asmbs.org/?p=3737.

An analytical tool has emerged that allows different quality “signals”, including reliability and risk adjusted volume, risk adjusted mortality and risk adjusted grade 2 and 3 complications to be combined into a single COMPOSITE measure of quality.\(^{(26)}\) The composite measure is unique in that it can predict with reasonable accuracy how a center will perform in the future based on its past performance. It provides multiple targets for quality improvement. Programs and surgeons can measure themselves against their peers around the nation and in their state. Coupled with regular feedback to programs/surgeons and process improvement, the composite measure can decrease morbidity and change practice behavior that will improve patient safety using evidence-based risk-adjusted results.\(^{(27)}\)

The data in Figure 1 illustrates the strength of the predictive value in bariatric surgery by composite measures. The composite measure explained the variability in comparing the rankings of hospitals to their subsequent performance. Hospital volume was the worst predictor.
Accreditation aside, the most important aspect of receiving data from the ASMBS quality program would be to engage in process improvement. Clinically rich data used to build a composite quality measure provides ample targets for process improvement. In one process improvement project carried out by the MBSC, a variation in the use of IVC filters was noted from 0-34 percent within the 20 hospitals participating in the collaborative at that time. Further analysis showed that patients with a filter placed preoperatively had a similar rate of VTE, serious complications including death, as patients without a filter. In addition, IVC filters were associated with a high degree of death and complications related to the filter itself. There were no characteristics of any patient in whom placement of a filter improved outcomes. As part of the process improvement initiative throughout the subsequent year, all programs were asked to change their practice patterns by eliminating IVC filter placement. Patient safety was shown to improve significantly, and there was a cost savings of approximately $2.6 million that year. As a result, the cost savings of this single intervention paid for the administrative costs of the Michigan collaborative program. The impact of process improvement is monumental in terms of improving patient safety and decreasing cost. This is especially true when patient safety and decreasing cost of interventions are important targets for healthcare reform.

In regards to measuring whether any given center is delivering high quality outcomes in bariatric surgery, it appears that using a composite measure provides the best tool for the ASMBS quality program to use in accreditation. The composite index of quality can be constructed whereby individual surgeons and programs can be given their data so that they can compare their overall performance in relationship to their peers. Eventually, with consent of the surgeons and hospitals, the composite could potentially be used for public reporting in 2014 and beyond as required by the public for transparency.

Although it would be possible to use some percentile of the index to serve as a discriminator in allowing programs to become part of the ASMBS quality program, another approach would allow all programs doing minimum hospital volume to engage in the program with a goal of having 100 percent of bariatric surgeons and hospitals performing bariatric procedures participating. The advantages are numerous, but include the ability of surgeons and new fellows to

<table>
<thead>
<tr>
<th>Odds Ratio (95% CI), 1-star vs. 3-star</th>
<th>Hospital volume</th>
<th>Serious complications</th>
<th>Composite measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.85 (0.43-1.68)</td>
<td>1.56 (0.84-2.91)</td>
<td>1.99 (1.14 -3.47)</td>
<td></td>
</tr>
<tr>
<td>% Variation Explained</td>
<td>0%</td>
<td>28%</td>
<td>89%</td>
</tr>
</tbody>
</table>
establish a program and understand and participate in a matrix of quality that begins at the beginning of their experience, instead of excluding them until they achieve volume. Another advantage is that we could potentially include populations that are currently excluded by the volume criteria like pediatric and adolescent programs as well as including all the bariatric surgeons/programs that do not have the local resources to mount an expensive tertiary bariatric effort.

How would the ASMBS/ACS National Quality Improvement Program work?

Programs would enter based on a new quality matrix that will be presented in the next segment. Programs would utilize a database that was able to generate composite quality measures and risk adjusted outcomes so that programs would understand where they were in relationship to the outcome measure for ongoing accreditation. All programs would participate at least twice per year in regional or national meetings on quality where specific targets for improvement would be presented and discussed. The state chapters could be engaged in establishing state and regional quality collaborative similar to Michigan and in conjunction with the ACS chapters where desirable. The society could provide the architecture, central administration and direction for the collaborative model.

1. Based on approval of the ASMBS Executive Council in September, the ASMBS BSCOE database (BOLD) is being used to construct a composite risk and reliability adjusted composite measure in bariatric surgery. The database was closed on October 1, cleaned and sent to Dr. John Birkmeyer and Dr. Justin Dimick in Michigan who authored multiple papers on this topic. The same database has been offered to the ACS to combine with and compare to the results in the Bariatric NSQIP database. The results of these two analyses will inform the ASMBS regarding the data establish a composite and generate initial targets for the national continuous quality improvement program.
2. At all times, regardless of funding (should a national payor be interested in participating in funding the project), the data would remain confidential unless released by the program/surgeon. The control of the aggregate data would rest with the society through the process already established. A goal would be to have that data become an open access database so that interested authors could use the data.
3. There would be a “roll in” period during which the use of the composite would be made and programs given a chance to come into compliance with a performance standard. By 2013, all programs would be expected to perform within statistical parameters yet to be determined.
4. Programs/surgeons will have the opportunity to consult with an ASMBS CQI team in order to receive expert advice as needed to improve the safety and quality of their program.
5. Implementation of the new program would occur beginning in March with applications being taken in July for new programs to the network. Current programs would renew into the new program, and the BSCOE would gradually be phased out. Programs that wanted to roll in sooner would be eligible. If fees had already been paid, the program would not be required to pay again until their annual renewal date.

In order to understand how the new quality matrix would work, please look for the third segment in this conversation. Remember that comments are welcome, and a survey will be sent out in the coming days.

References

2. ACS Surgery: Principles and Practice Elements of Contemporary Practice Birkmeyer JD and Dimick JB. Chapter 2 Performance Measures in Surgical Practice