Measuring the Quality of Barrett's Esophagus Management With Measures That Are High Quality

In 2009, Americans underwent approximately 350,000 upper endoscopies for surveillance of Barrett's esophagus (BE), and underwent an additional 1,650,000 upper endoscopies for evaluation of symptoms of gastroesophageal reflux disease, which often is done to assess for BE.[1] The total societal cost of this endoscopic practice was an estimated $3.5 billion.[1]

To maximize the utility of upper endoscopy as a preventive tool for esophageal cancer and justify the value of the procedure, it is important to assure that patients are receiving a high-quality examination. Unfortunately, evidence suggests that there is significant variability in the concordance of endoscopic procedures with published guidelines. For instance, in a database of a large national pathology practice, only 51% of endoscopies provided 4 specimens every 2 cm of the length of BE.[2] Among Medicare recipients with BE, 59% underwent a repeat upper endoscopy within 3 years of their previous procedure, rather than the recommend interval of 3–5 years.[3] Furthermore, screening and surveillance for esophageal adenocarcinoma may fall short in other domains as well. For example, in a sample of patients with BE in North Carolina, 68% of patients overestimated their risk of developing esophageal cancer, suggesting a failure of communication with their endoscopists.[4] Most important, in multiple studies, <10% of patients with esophageal adenocarcinoma had been diagnosed with BE before their presentation with cancer.[5,6]

To improve the value of endoscopies performed for screening and surveillance of BE, we first need to measure reliably endoscopist performance. Prior gastroenterology society guidelines and expert panels have provided recommendations on the management of BE, but few have explicitly established how the quality of that management should be measured.[7–11] Thus, a group of experts in BE met during the American Gastroenterological Association (AGA) Freston Conference in Chicago of August 2013, aiming to reach consensus on a list of quality measures for management of BE.[12] These experts used a modified Delphi approach with a threshold of 80% agreement to determine acceptance on the final list of quality measures. Although this threshold is within the range reported in other Delphi consensus studies,[13] the end result is that this panel chose to incorporate measures that some experts actually disagreed with. It is difficult to imagine holding endoscopists to a practice quality measure where even experts cannot find agreement on, and thus one could argue that the agreement threshold should have been higher. Unfortunately, the authors did not conduct a systematic review of the literature, nor were the panelists presented with the relevant articles before voting. Although the experts were likely familiar with much of the literature regarding each statement, each individual expert is unlikely to have been fully aware of all the details of every relevant article for each statement. At the conclusion of the process, the authors agreed on 8 statements that were each felt to be supported by moderate quality evidence.[12]

The Freston initiative is important because it reinforces the importance of quality measurement and quality improvement in the contemporary healthcare landscape. We applaud the efforts of participants in taking ownership of quality as a key factor in the optimal management of BE. Indeed, the Department of Health and Human Services (HHS) has laid out ambitious goals for reforming health care delivery through increased use of incentives to motivate higher value care. For example, HHS has a goal to link 90% of all Medicare fee-for-service payments to quality or value by 2018.[14] For gastroenterology practices to thrive in this value-based marketplace, we need to be at the forefront of quality measure development, rather than having ill-conceived measures foisted on us.

Whereas the effort of the Freston initiative is commendable, there are critical deficiencies in the proposed quality metrics that may limit their successful implementation. These deficiencies are perhaps best highlighted by evaluating the proposed measures in the context of the National Quality Forum's criteria for high quality measures (Table 1).[15] Although this framework is designed to evaluate fully specified measures, it offers a useful guide for measure concept evaluation. The framework has five criteria. The first criterion relates to the importance of the measure: is the measure evidence based and does it address a priority performance gap? The second criterion relates to the reliability and validity of the measure: is the measure well-specified to allow it to be compared across organizations? Does the measure adequately differentiate high and low quality clinicians? The third criterion relates to the feasibility of the measure: are the data required for the measure readily available or easily captured? The fourth criterion refers to the usability of the measure: can the measures be used by stakeholders in both accountability and quality improvement programs to improve the quality of care? Finally, the last criterion evaluates for competing measures: are there other measures for the same concept?

Based on these criteria, most of the proposed quality measures in this article have substantial deficiencies (Table 1). For instance, the measures regarding documentation of endoscopic landmarks (#1 and #2) have not been documented to differentiate the quality of care in terms of patient-relevant outcomes, such as cancer diagnosis. Such measures would not be particularly useful for measuring accountability and directing performance improvement. Nonetheless, reporting the length of BE using Prague classification would be an important component of a high-quality report to implement a measure of adequacy of biopsies obtained. For instance, fewer biopsies should be

Table 1. Assessment of Proposed BE Quality Measures Based on the National Quality Forum's Criteria for High Quality Measures

Statements agreed upon by Freston experts	Performance gap has been documented	Measure is well specified	Differentiates high- from low-quality outcomes	Feasible to capture without undue burden	Usable for accountability and performance improvement	No competing measures
The squamocolumnar junction, the gastroesophageal junction, and the location of the diaphragmatic hiatus should be recorded.	-	-	-	+	-	+
The endoscopist should document the extent of suspected BE using Prague criteria.	-	-	-	+	-	+
The normal-appearing and normally located squamocolumnar junction should not be biopsied.	-	-	+	+	+	+
In a BE patient without dysplasia, follow-up surveillance endoscopy should be no sooner than in 3–5 years.	+	-	-/+	-/+	+	+
In a BE patient undergoing surveillance endoscopy, systematic biopsies should be taken every 1–2 cm in 4 quadrants throughout the extent of the endoscopically involved segment.	+	-	+	-/+	+	+
In a BE patient undergoing surveillance endoscopy, biopsies from any visible raised or depressed lesions should be obtained and processed separately from the systematic biopsies.	-	-	-/+	+	+	+
In patients with dysplastic BE or early esophageal adenocarcinoma, a diagnostic endoscopic resection should be performed on any raised or suspicious areas.	-	-	-/+	-/+	+	+
In patients with BE-associated neoplasia, the goal of the endoscopic treatment should be complete eradication of the BE segment in addition to any dysplastic lesions.	-	-	+	-	-	+

BE, Barrett's esophagus; +, proposed measure meets the criterion; -, proposed measure does not meet the criterion; -/+, equivocal whether it meets the criterion.