

Quality measures in endoscopy: A systematic analysis of the overall scientific level of evidence and conflicts of interest



Authors

Simcha Weissman¹, Muhammad Aziz², Matthew R. Baniqued³, Vikas Taneja⁴, Mohammed El-Dallal⁴, Wade Lee-Smith², Sameh Elias¹, Joseph D. Feuerstein⁵

Institutions

- 1 Department of Medicine, Hackensack Meridian Health Palisades Medical Center, North Bergen, New Jersey, United States
- 2 Division of Gastroenterology and Hepatology, University of Toledo Medical Center, Toledo, Ohio, United States
- 3 Hackensack Meridian Health School of Medicine, Hackensack, New Jersey, United States
- 4 Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, United States
- 5 Division of Gastroenterology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, United States

submitted 22.7.2021

accepted after revision 15.3.2022

Bibliography

Endosc Int Open 2022; 10: E776–E786

DOI 10.1055/a-1809-4219

ISSN 2364-3722

© 2022. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Corresponding author

Joseph D. Feuerstein MD, Associate Clinical Chief,
Gastroenterology, Center For Inflammatory Bowel Disease,
Beth Israel Deaconess Medical Center, Associate Professor of
Medicine Harvard Medical School, 330 Brookline Ave, Boston,
MA 02215

Fax: +1-617-632-9199

jfeuerst@bidmc.harvard.edu

ABSTRACT

Background and study aims Quality measures were established to develop standards to help assess quality of care, yet variation in endoscopy exists. We performed a systematic review to assess the overall quality of evidence cited in formulating quality measures in endoscopy.

Methods A systematic search was performed on multiple databases from inception until November 15, 2020, to examine the quality measures proposed by all major societies. Quality measures were assessed for their level of quality evidence and categorized as category A (guideline-based), category B (observational studies) or category C (expert opinion). They were also examined for the type of measure (process, structure, outcome), the quality, measurability, review, existing conflicts of interest (COI), and patient participation of the quality measure.

Results An aggregate total of 214 quality measures from nine societies (15 manuscripts) were included and analyzed. Of quality measures in endoscopy, 71.5%, 23.8%, and 4.7% were based on low, moderate, and high quality of evidence, respectively. The proportion of high-quality evidence across societies was significantly different ($P=0.028$). Of quality measures, 76% were quantifiable, 18% contained patient-centric outcomes, and 7% reported outcome measures. None of the organizations reported on patient involvement or external review, six disclosed existing COI, and 40% were published more than 5 years ago.

Conclusions Quality measures are important to standardize clinical practice. Because over 70% of quality measures in endoscopy are based on low-quality evidence, further studies are needed to improve the overall quality to effectively set a standard, reduce variation, and improve care in endoscopic practice.

Introduction

Quality measures are used to provide a standardized metric by which the overall quality of healthcare being offered to patients can be assessed [1]. In general, they can be used to assess characteristics of care (structural), the delivery of care (process), or the results of care (outcomes). In particular, these measures can be developed in relation to diagnostics, management, patient prevention, or administration function [2]. While quality measures can be used as a means to identify those providing high-quality care and thus provide a mechanism to reward those for this practice, they can also be used as a means to penalize those who fail to meet the expected standards [3].

Owing to the variation in colorectal cancer screening recommendations, lower gastrointestinal endoscopy was one of the first areas of endoscopy to directly address quality [2–6]. As such, numerous potential measures of quality in lower endoscopy have been identified, and as a consequence, many professional societies have published recommendations on performance measures [2–4]. The united aim was to propose quality and safety procedures and indicators to facilitate quality improvement in digestive endoscopy units [2–5]. However, these recommendations are country-specific and not always evidence-based, which has limited their wider adoption [2, 4, 7]. Hence, while the goal of guidelines and position statements are to reduce variation in practice and standards between individual endoscopists and centers, data assessing the quality of evidence supporting quality measures are lacking [1, 8–10].

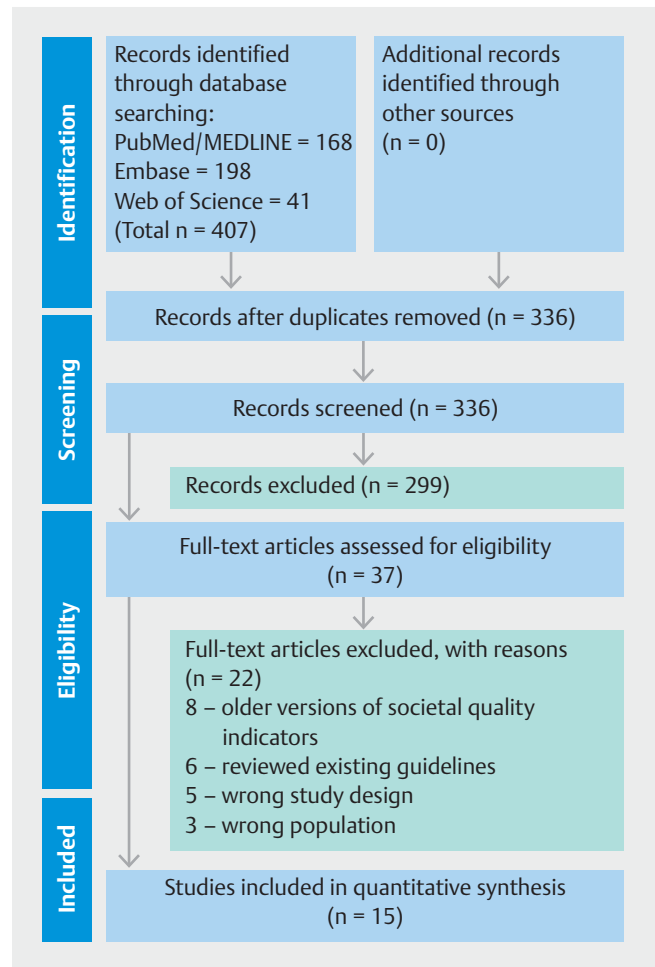
Despite attempts at using quality measures to standardize healthcare, significant variation in clinical practice remains [11]. When evaluating the reasons for non-adherence to guideline recommendations, some report a lack of confidence in guidelines due to the lack of high-quality evidence supporting many of the recommendations [11]. Similarly, the ability for quality metrics to effectuate change in clinical practice, standardize care, and improve the quality of care when they are based on lower-quality evidence is not proven to provide improved patient outcomes in long-term longitudinal studies [1, 8–13].

We, therefore, conducted a systematic review of the quality measures in endoscopy proposed by international medicine, oncology, surgical, gastrointestinal, and endoscopy societies to assess the overall quality of evidence and COI cited in formulating these quality measures.

Methods

Search strategy and data

A systematic literature search was performed on the PubMed/MEDLINE, EMBASE, and Web of Science databases using Mesh terms; “endoscopy,” “digestive endoscopy,” “gastrointestinal endoscopy,” “quality standards,” “quality measures,” “quality indicators,” and “quality metrics” in different combinations to generate a comprehensive and up-to-date list of articles on November 15, 2020. In addition, major international medicine, gastrointestinal, and endoscopy society websites were also examined for the presence of endoscopy-specific quality meas-



► **Fig. 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement diagram delineating the process by which manuscripts were screened and ultimately included.

ures. Moreover, in all manuscripts identified, citations were examined for relevant papers. This identified a total of 407 manuscripts. After screening for relevance and excluding studies that: 1) reviewed quality measures; 2) only discussed adherence to quality measures; and/or 3) did not discuss the presence of endoscopy-specific quality measures, 15 manuscripts, totaling nine task forces/groups of societies remained and were included in the final analysis. Manuscripts were not limited by age, date, or language written. All studies were screened by two authors (SW and MB) and any disagreement was resolved by mutual discussion and by consulting a third author (JDF) via a modified Delphi system [14]. The methodological protocol herein was established a priori as we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) to conduct our systematic review (► **Fig. 1**) [15, 16].

Quality measures

All quality measures were examined for: 1) the type of measure—structural, process, or outcome related; 2) the grading methods used; 3) the supporting quality of evidence behind the in-

clusion of the quality measure; 4) whether the quality measure can be numerically measured; 5) if the measure was externally reviewed; 6) if there was inclusion of patients in the development of the measure; 7) if the measure reported the presence of any conflicts of interest (COI); 8) if the measure could impact patient outcomes; and 9) its age from publication.

Levels of evidence

Given the diversity of grading systems internationally, we formulated levels of evidence based on the GRADE and ABC(D) models of level of evidence used in the development of clinical practice guidelines and prior studies assessing guideline quality [1, 8–10, 12, 13].

- Category A: High-quality of evidence: Based on clinical guidelines derived from randomized controlled trials
- Category B: Moderate-quality of evidence: Based primarily on observational, population-based, or cross-sectional studies
- Category C: Low-quality of evidence: Based primarily on expert opinion or small case-series with weak evidence or high study heterogeneity.

When evidence was based on prior studies (moderate quality of evidence), these studies were analyzed for the methodology used. If the quality measure had no accompanying primary literature cited, it was subsequently placed in the category of low-quality evidence (i.e. expert opinion), as done in prior guideline quality studies [1, 5].

Review of the quality measures

All quality measures were reviewed by two authors (SW and MB) for determination of the type of measure, the supporting quality of evidence behind the measure, whether in fact it can be measured, if it was externally reviewed, if patients were included in its development, if it reported the presence of any COI, and if the measure could impact patient outcomes. COI that were determined to be relevant included being a part of an advisory board, speaker's bureau, and consulting or industry-sponsored continuing medical education activities (government and non-profit awards were not considered COI). If there was disagreement between the above authors with regards to data extraction, a third author (JDF) reviewed it using a modified Delphi system [14].

Ethical considerations

Given the publicly available nature of these data, i.e. all recommendations were previously published and patients were not individually included, it is exempt from Institutional Review Board review. In addition, informed consent was not needed as these data were not obtained from study participants.

Data analysis

Quality measures were assessed for evidence quality and categorized as category A (guideline-based), category B (primarily observational/population-based studies), or category C (expert opinion). Statistical analysis was conducted in R using ANOVA,

linear regression and chi-square or Kruskal-Wallis tests. A $P=0.05$ was considered significant.

Results

Organizations involved and quality measures reported

The following nine (task forces/groups of) societies/organizations quality measures (comprising 15 manuscripts) were included in the final analysis: American Gastroenterological Association (AGA) [17], American College of Gastroenterology and American Society of Gastrointestinal Endoscopy (ACG-ASGE) [18–23], British Society of Gastroenterology and Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland (BSG-AUGIS) [24], Canadian Association of Gastroenterology (CAG) [25], European Society of Gastrointestinal Endoscopy (ESGE) [26, 27], Health Programme of the European Union (HPEU) [28], National Colorectal Cancer Roundtable (NCCR) [29], Sociedad Española de Patología Digestiva (SEPD) [30], and the Spanish Society of Gastroenterology and Spanish Society of Gastrointestinal Endoscopy Working Group (SSG-SSGE) [31].

A total of 183 distinct and an aggregate total of 214 quality measure recommendations were reviewed and included in this study from the 15 manuscripts as reported by the nine task forces/groups of societies/organizations: AGA reported 7, ACG-ASGE reported 36 quality measures, BSG-AUGIS reported 38, CAG reported 23, ESGE reported 44, HPEU reported 29, NCCR reported 4, SEPD reported 13, and SSG-SSGE reported 20 (► **Table 1**) [17–31].

Types of measures

Fifty-three percent of quality measures reported process measures, 40% reported structure measures, and 7% reported outcome measures. The AGA only reported process-based measures. Sixty-nine percent of ACG-ASGE quality measures were process measures and 31% were structure measures. Sixty-eight percent of BSG-AUGIS quality measures were process measures and 32% were structure measures. 26% of CAG quality measures were process measures and 74% were structure measures. Forty-five percent of ESGE quality measures were process measures, 45% were structure measures, and 10% were outcome measures. 52% of HPEU quality measures were process measures and 48% were structure measures. Seventy-five percent of NCCR quality measures were process measures and 25% were structure measures. Sixty-nine percent of SEPD quality measures were process measures, 8% were structure measures, and 23% were outcome measures. Fifteen percent of SSG-SSGE quality measures were process measures, 50% were structure measures, and 35% were outcome measures [17–31].

Grading method used

Four organizations (BSG-AUGIS, CAG, ESGE, and SEPD) used the Grading of Recommendation Assessment Development and Evaluation (GRADE) system, one (SSG-SSGE) used the Center for Evidence Based Medicine (CEBM) from Oxford method, and

► **Table 1** Distinct quality measure recommendations listed by level of scientific evidence.

High-quality evidence	Type	Patient-centric	Quantifiable
Exams should be performed only after adequate bowel preparation i. e. without any residual stool or liquid in the lumen that could mask any suspicious area	Process	Yes	Yes
Exams should be complete to the caecum and there should be slow, careful inspection of the colonic mucosa during withdrawal of the scope	Process	Yes	Yes
Where coeliac disease is suspected, a minimum of four biopsies should be taken, including representative specimens from the second part of the duodenum and at least one from the duodenal bulb	Process	No	Yes
Attention should be focused on preventing transmission of highly resistant organisms by duodenoscopes, in particular, on ensuring cleaning and HLD of the elevator mechanism and elevator wire channel	Process	No	No
General infection control principles should be complied with in the endoscopy unit	Structural	No	No
Use of standard precautions reduces the transmission of infection from patients to endoscopy personnel	Process	Yes	Yes
Adenoma detection rate	Outcome	Yes	Yes
Adenoma detection rate after positive FOBT	Outcome	No	Yes
Appropriate interval between colonoscopies	Process	Yes	Yes
Mean number of adenomas excised per colonoscopy procedure	Structural	No	Yes
Moderate-quality evidence			
Endoscopy facilities should ensure that the services they provide are patient-centered	Structural	No	No
Patients can be divided into low, intermediate and high-risk groups with respect to their risk of developing advanced adenomas and cancer based on findings at baseline colonoscopy. The surveillance strategy can vary accordingly	Process	Yes	Yes
A readjustment of the strategy can be made based on findings at the first and subsequent surveillance examinations	Structural	Yes	No
Low risk. Patients with only one or two small (< 10 mm) adenomas are at low risk, and should be returned to the screening program	Process	No	No
Intermediate risk. Patients with three or four small adenomas or at least one adenoma of size ≥ 10 mm and < 20 mm are at intermediate risk	Process	Yes	Yes
Intermediate risk. Patients with three or four small adenomas or at least one adenoma of size ≥ 10 mm and < 20 mm are at intermediate risk and should be offered surveillance at 3-yearly intervals	Process	Yes	Yes
High risk. If either of the following is detected at any single examination (at baseline or follow-up): 5 or more adenomas, or an adenoma ≥ 20 mm, the patient is at high risk and an extra examination should be undertaken within 12 months, to check for missed synchronous lesions, before initiating 3-yearly surveillance. In the absence of evidence on the safety of stopping surveillance in the high-risk group, surveillance should continue	Process	Yes	Yes
High risk. If either of the following is detected at any single examination (at baseline or follow-up): 5 or more adenomas, or an adenoma ≥ 20 mm, after two consecutive normal exams, the interval can be extended to 5-yearly	Process	Yes	Yes
The risk stratification is based on accurate detection and complete removal of adenomas otherwise risk status will be underestimated	Process	No	No
Recommendations should not differ for patients with a family history who are found to have adenomas, unless it is suspected that they have one of the dominantly inherited conditions	Structural	No	No
New symptoms should be assessed on the basis that a recent clearance colonoscopy reduces the chance of advanced adenomas and cancers but does not eliminate the risk altogether	Process	No	No
By their nature locally removed pT1 cancers are high-risk lesions and therefore should undergo a surveillance strategy similar to the high-risk adenoma group	Structural	No	No
There is no evidence that patients in whom only small, distally located hyperplastic polyps are detected are at increased risk for colorectal cancer; therefore, they should be offered routine screening	Structural	No	No
A safety checklist should be completed before starting an esophagogastroduodenoscopy (EGD)	Process	No	Yes
Intravenous sedation and local anesthetic throat spray can be used in conjunction if required. Caution should be exercised in those at risk of aspiration	Process	No	No

► **Table 1** (Continuation)

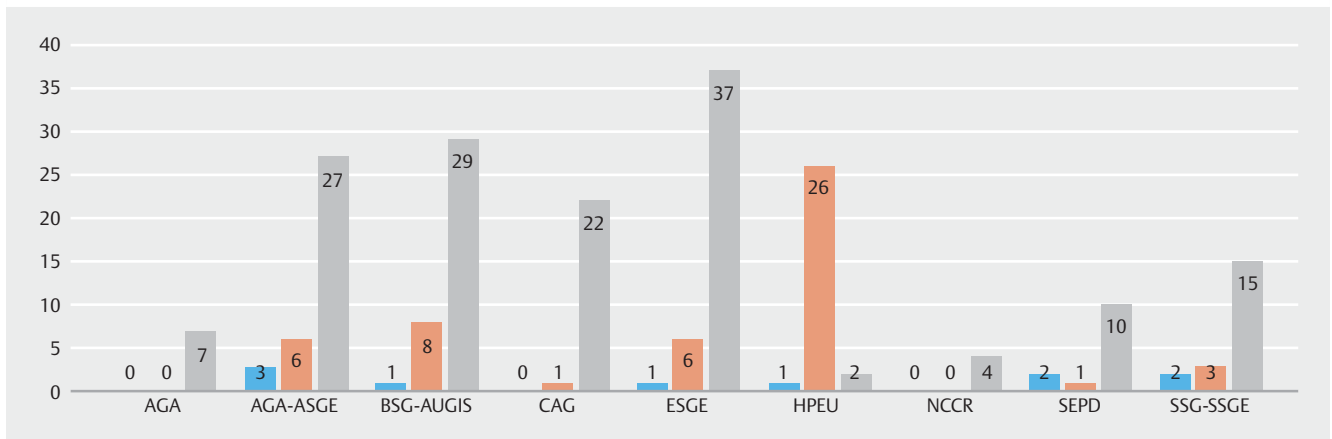
High-quality evidence	Type	Patient-centric	Quantifiable
Adequate mucosal visualization should be achieved by a combination of adequate air insufflation, aspiration and the use of mucosal cleansing techniques	Process	No	No
When no lesions are detected within a Barrett's segment, biopsies should be taken in accordance with the Seattle protocol	Process	No	Yes
If squamous neoplasia is suspected, full assessment with enhanced imaging and/or Lugol's chromoendoscopy is required	Process	No	Yes
Biopsies from two different regions in the esophagus should be taken to rule out eosinophilic esophagitis in those presenting with dysphagia/food bolus obstruction, where an alternate cause is not found	Process	No	Yes
Where gastric or duodenal ulcers are identified, <i>H. pylori</i> should be tested and eradicated if positive	Process	No	Yes
The presence of gastric polyps should be recorded, with the number, size, location and morphology described, and representative biopsies taken	Process	No	Yes
In the event of reprocessing failure, the patient, the institution's designated infection control personnel, local and/or state public health agencies, the FDA, the CDC, and the manufacturers of the involved equipment should be notified immediately	Structural	No	No
Frequency with which endoscopy is performed for an indication that is included in a published standard list of appropriate indications, and the indication is documented (priority indicator)	Process	No	Yes
Colonoscopy withdrawal time	Process	No	Yes
Cecal intubation rate	Process	No	Yes
Decontamination indicators	Structural	No	No
We recommend endoscopy services have policies and processes in place to assess the appropriateness of procedures against guidelines and take action when endoscopic procedures have been performed inappropriately	Structural	No	No
We recommend endoscopy services have procedures in place to assess the comfort of patients before, during, and after procedures	Structural	No	No
We recommend that endoscopy services provide an environment and have processes in place that ensure the privacy and dignity of patients is respected and maintained	Structural	No	No
Endoscopes should undergo HLD as recommended by governmental agencies and all pertinent professional organizations for the reprocessing of gastrointestinal endoscopes	Structural	No	Yes
The efficacy of manual cleaning and HLD is operator dependent, thus assignment of personnel responsible for endoscope reprocessing, extensive training of reprocessing personnel, process validation, and quality assurance is vital, and staff competency should be assessed at the very least on an annual basis	Structural	No	Yes
Low-quality evidence			
For a patient to give a physician informed consent to perform an elective endoscopic procedure, the patient must be advised, in a timely fashion, of all relevant information about the procedure, its risks, benefits and alternatives, if any, and be given an opportunity to ask questions that the physician must answer	Process	No	No
Endoscopy facilities should meet or exceed defined operating standards, in all domains, consistent with accreditation under the appropriate national or regional standards	Structural	No	Yes
Endoscopic procedures are performed for an appropriate, clearly documented indication, consistent with current, evidence-based guidelines	Process	No	Yes
Endoscopy facilities should have the technical and personnel resources required by national and/or regional standards to complete all planned procedures safely and effectively	Structural	No	Yes
Endoscopy facilities should implement and monitor the effect of pre-procedure policies that ensure best practice	Structural	No	No
Endoscopy facilities should implement and monitor the effect of intraprocedural policies that ensure best practice	Structural	No	No
The endoscopy facility should implement and monitor the effects of policies for the discharge of patients that ensure best practice	Structural	Yes	Yes

► **Table 1** (Continuation)

High-quality evidence	Type	Patient-centric	Quantifiable
Endoscopy facilities should ensure that there is a policy in place to notify patients of the need, and appropriate interval, for follow-up	Structural	Yes	No
All patients, on discharge, are given written information regarding the procedural findings, plans for treatment and follow-up, worrisome symptoms to watch for, and steps to be taken	Process	Yes	Yes
Endoscopy facilities should maintain a comprehensive quality improvement program incorporating formal, regular, scheduled review of performance reports	Structural	No	No
Endoscopy facilities should appoint a review committee to monitor and report back to management on adherence to and implementation of quality standards	Structural	No	No
Endoscopy facilities should systematically and regularly review current indicators of quality for all endoscopic procedures and implement appropriate responses	Structural	No	No
Endoscopy facilities should systematically and regularly review current indicators of safety for all endoscopic procedures and implement appropriate responses	Structural	No	No
Endoscopy facilities should provide high-quality education programs or opportunities for all staff	Structural	No	No
All endoscopy facility personnel in training should be supervised and their performance monitored regularly until they have achieved competency to perform specified routine and/or emergency procedures according to appropriate current standards	Process	No	Yes
All endoscopy facility personnel engaged, directly or indirectly, in endoscopy service delivery should be trained and certified as having competency to perform specified routine and/or emergency procedures according to appropriate current standards	Structural	No	Yes
Endoscopists should regularly review their endoscopic practice and outcome data with the aim of continuous professional development	Structural	No	No
Endoscopists should be granted privileges to perform specified procedures based on a formal evaluation of their competence consistent with appropriate current standards	Process	No	No
Endoscopists' privileges should be subject to formal, regular, scheduled review to ensure that renewal is based on documented competence to perform specified procedures consistent with appropriate current standards	Process	No	No
Endoscopic procedures should be reported in a standardized electronic format, including mandatory reporting fields, to provide full documentation of all necessary clinical and quality measures	Structural	No	No
Endoscopy facilities should implement policies to monitor and ensure the timeliness and completeness of procedure reporting	Structural	No	No
Endoscopy facilities should systematically and at least annually solicit patient feedback, report the results to the service and to the institution's quality committee, and implement effective measures to address patients' concerns	Structural	Yes	No
Intermediate risk. Patients with three or four small adenomas or at least one adenoma of size ≥ 10 mm and < 20 mm, after one negative exam, the interval for surveillance can be extended to 5 years. After two consecutive normal exams, the patient can return to routine screening	Process	Yes	Yes
High risk. If either of the following is detected at any single examination (at baseline or follow-up): 5 or more adenomas, or an adenoma ≥ 20 mm, in the absence of evidence on the safety of stopping surveillance in the high-risk group, surveillance should continue	Process	No	Yes
Patients with a failed colonoscopy should, if possible, undergo repeat colonoscopy or an alternative complete colonic examination, particularly if they are in the high-risk group	Process	Yes	No
The site of large sessile lesions removed piecemeal should be re-examined at 2 to 3 months. Small areas of residual tissue can then be treated endoscopically, with a further check for complete eradication within 3 months. India ink tattooing aids recognition of the site of excision at follow-up. If extensive residual lesion is seen, surgical resection must be considered, or alternatively, referral to a colonoscopist with special expertise in advanced endoscopic excision	Process	Yes	No
The decision to undertake each colonoscopic surveillance examination should depend not only on adenoma characteristics, but also on the patient's age and wishes, and the presence of significant comorbidity. The patient status should be established prior to attendance for each examination	Structural	No	No
The cut-off age for stopping surveillance is usually 75 years, but this should also depend upon patient wishes and comorbidity	Structural	No	No

► **Table 1** (Continuation)

High-quality evidence	Type	Patient-centric	Quantifiable
Following cessation of surveillance, individuals should be returned to the population screening program	Structural	No	No
The potential benefit of supplementing colonoscopy exams with fecal occult blood testing is presumed to be too small to warrant double testing; therefore, it is recommended to stop fecal occult blood testing in individuals who are undergoing surveillance	Structural	No	No
For surveillance purposes, serrated adenomas (traditional serrated adenomas and mixed polyps with at least one adenomatous component) should be dealt with like any other adenoma; there are no data to suggest that different surveillance intervals are required	Structural	No	No
One or more large (≥ 10 mm) hyperplastic polyps or other non-neoplastic serrated lesions anywhere in the colon or multiple smaller lesions of these types in the proximal colon may confer an increased risk, but there are no data available to indicate appropriate surveillance intervals	Process	No	No
Every screening program should have a policy on surveillance. The policy may limit surveillance to the high-risk group if sufficient resources are not available to include people with lower risk	Structural	No	No
The responsibility of program management to assure the quality of screening services includes quality assurance of surveillance. For surveillance, the same principles, methods and standards of quality assurance apply that are elucidated elsewhere in the first edition of the European Guidelines	Structural	No	No
Adherence to the guidelines should be monitored	Structural	No	No
Surveillance histories should be documented and the results should be available for quality assurance	Structural	No	Yes
The occurrence of colorectal cancer in any individual in whom adenomas or pT1 cancers have been detected at a previous exam should be captured as an auditable outcome for any surveillance program	Structural	No	No
Patients should be assessed for fitness to undergo a diagnostic EGD	Process	No	Yes
Patients should receive appropriate information about the procedure before undergoing an EGD	Process	No	No
An appropriate time slot should be allocated dependent on procedure indications and patient characteristics	Structural	No	No
A checklist should be undertaken after completing an EGD, before the patient leaves the room	Process	No	Yes
Only an endoscopist with appropriate training and the relevant competencies should independently perform EGD	Structural	No	No
We suggest that endoscopists should aim to perform a minimum of 100 EGDs a year, to maintain a high-quality examination standard	Structural	No	Yes
UGI endoscopy should be performed with high-definition video endoscopy systems, with the ability to capture images and take biopsies	Structural	No	No
A complete EGD should assess all relevant anatomical landmarks and high-risk stations	Process	No	Yes
Photo documentation should be made of relevant anatomical landmarks and any detected lesions	Process	No	Yes
The quality of mucosal visualization should be reported	Process	No	Yes
It is suggested that the inspection time during a diagnostic EGD should be recorded for surveillance procedures, such as Barrett's esophagus and gastric atrophy/intestinal metaplasia surveillance	Structural	No	Yes
Where a lesion is identified, this should be described using the Paris classification and targeted biopsies taken	Process	No	Yes
Endoscopy units should adhere to safe sedation practice	Structural	No	No
The length of a Barrett's segment should be classified according to the Prague classification	Process	No	Yes
Where a lesion is identified within a Barrett's segment, this should be described using the Paris classification and targeted biopsies taken	Process	No	Yes
Esophageal ulcers and esophagitis that is grade D or atypical in appearance, should be biopsied, with further evaluation in 6 weeks after PPI therapy	Process	Yes	Yes
The presence of an inlet patch should be photo-documented	Structural	No	Yes
The presence of a hiatus hernia should be documented and measured	Structural	No	Yes
Varices should be described according to a standardized classification	Process	No	Yes



► **Fig. 2** Clustered column chart depicting the graded evidence quality by society/organization. Blue represents high-quality evidence, orange represents moderate-quality evidence, and gray represents low-quality evidence.

four (AGA, ACG-ASGE, HPEU and NCCR) did not use a strict methodology or created their own methodology to formulate/provide evidence regarding their quality measures [17–31].

Quality of levels of evidence

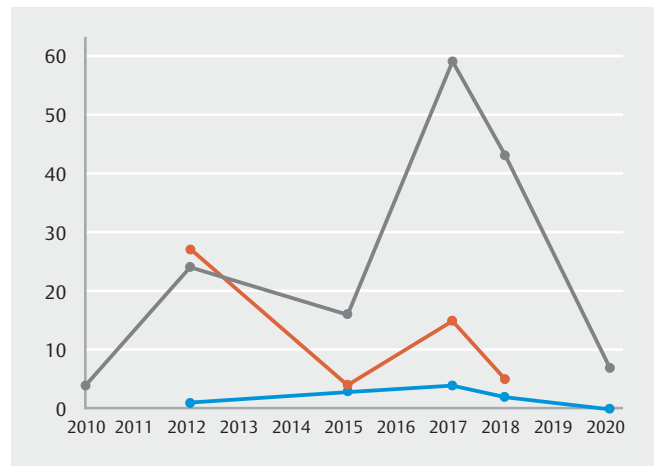
An aggregate total of 214 quality measure recommendations were analyzed for their quality of level of evidence. Of quality measures, 4.7% (10) were category A, 23.8% (51) were category B, and 71.5% (153) were category C. Of these, there was disagreement among the two data extracting authors (SW and MB) regarding the level of evidence pertaining to 27 quality measures—of which 19 were resolved via mutual discussion, and the remaining eight by a third author (JDF). The breakdown by society on the quality of level of evidence is shown in ► **Fig. 2**. The proportion of high-quality evidence across societies was significantly different ($P=0.028$) [17–31].

Measurability

Fifty-seven percent of recommendations reported measurable/quantifiable outcomes and 43% reported non-quantifiable outcomes. Ninety percent of category A quality measures recommendations were quantifiable, 75% of category B quality measures were quantifiable, and 54% of category C quality measures were quantifiable. Seventy-five percent of ASGE quality measures, 76% of BSG-AUGIS, 30% of CAG, 32% of ESGE, 31% of HPEU, 75% of NCCR, 100% of SEPD, and 65% of SSG-SSGE quality measures were quantifiable, respectively [17–31].

External review, patient participation, COI, and evidence cited

None of the nine organizations reported external review of their quality measures or included patients in the development of their quality measures. Six organizations (AGA, ACG-ASGE, CAG, ESGE, HPEU, and NCCR) reported the presence of a COI when it existed, whereas the remaining three did not. Five organizations (ACG-ASGE, ESGE, HPEU, SEPD and SSG-SSGE) cited evidence behind the grade assigned for their quality measures, whereas the remaining four did not [17–31].



► **Fig. 3** Scatter plot diagram depicting the graded evidence quality by the year the measure was published. Blue represents high-quality evidence, orange represents moderate-quality evidence, and gray represents low-quality evidence.

Effect on patient outcomes

Only 18% of all quality measures were directed toward improving patient outcomes. None of AGA quality measures were directed toward improving patient outcomes. Three percent of ACG-ASGE, 8% of BSG-AUGIS, 17% of CAG, 11% of ESGE, 38% of HPEU, 50% of NCCR, 46% of SEPD, and 30% of SSG-SSGE quality measures were patient outcome centric, respectively. Among process and outcome-based quality measures, 19% and 78% led to patient outcomes, respectively. Only 6% of structure-based quality measures were directed toward patient outcomes [17–31].

Age of publication

Forty percent of quality metrics were published more than 5 years ago (between 2010 and 2015), and 73% were published more than 3 years ago (between 2010 and 2017). There were no significant associations between publication year and evi-

dence quality ($P=0.17$). The distribution of evidence quality by publication year is represented in ► **Fig. 3**.

Discussion

Our study indicated that most (71.5%) of the current quality measures in endoscopy are based on lower-quality levels of evidence. Additionally, only about half (57%) of quality measure recommendations reported quantifiable outcomes, less than 10% reported outcome measures, and less than 20% of all quality measures were directed toward improving patient outcomes. Furthermore, many organizations did not include patients in the development of their quality measures, report on an external review of the guideline, include a strict grading system methodology, or report on the presence of existing COI. Finally, 40% of quality metrics were published more than 5 years ago and 73% more than 3 years ago.

With the publication of the Institution of Medicine (IOM) reports regarding quality of care in medicine, over the last decade, there has been a transformation toward a new found focus on the standardization of healthcare across different settings, including gastrointestinal endoscopy [32, 33]. Significant efforts have been made by different national and international societies to regulate quality measures for endoscopy units and physicians performing endoscopy [2–7]. The notion behind these efforts is to provide practitioners with a standard (benchmark) to track and compare actual performance. Despite these efforts, there is a significant disparity noted between the actual recommendations and the evidence behind these judgments as evident from our systematic review. One of the biggest challenges encountered is the fact that more than 70% of these metrics is centered around low-quality evidence.

Low-quality evidence creates substantial variation in the actual delivery of healthcare. These measures are based either on expert opinion or small studies with considerable heterogeneity. Given there is less science supporting these measures or metrics, practitioners may opt to discard these recommendations and instead opt for personal judgment and anecdotal evidence to cater the need for their patients. While the existing studies that led to formulation of these low-quality evidence cannot be changed, there is certainly a need to undertake high-quality studies that will allow societies to strengthen these measures. In this vein, the authors understand that it is the role of a society to cover all aspects of a technique or procedure, even if some aspect has not been adequately evaluated with high-level of evidence. Thus, while we call for higher-quality studies, the authors congratulate the various international societies for their work, as it is indeed challenging to produce a quality measure when there is a low-level of evidence.

Another area that raises concern is the lack of strict grading methods when formulating these measures and guidelines. Four major societies did not employ a standard grading methodology. In addition, three societies did not report on COI. None of the societies included patients while formulating these recommendations. While not specifically developed for quality metrics, societies should still follow IOM standards similar to guideline development. The development process should be

set a priori with a clear and transparent process that includes a standard methodology for grading evidence, reporting of all COI and how they will be handled, a process for external review of the manuscript, including a patient representative in the guideline panel, and report only quantifiable outcomes that are patient-centric [34]. The absence of current COI information among multiple guidelines is also notable. While the Institute of Medicine recommends that guideline panels should attempt to minimize COI, the disclosure of COI of panel members' is crucial to mitigate any potential undue industry influence and improve transparency.

Ideally quality metrics can be used to implement benchmarks as quality measures to enhance the performance of endoscopy and specifically to improve patient outcomes. The successful implementation should result in improved efficiency, reliability, and cost-effectiveness in the endoscopy unit. Quality measures should be mandated in some form to standardize care delivered to patients. The National Quality Forum (NQF), a public-private organization created in 1999, in response to the Advisory Commission on Consumer Protection and Quality in the Healthcare Industry, advises Centers for Medicare & Medicaid Services on the selection of performance measures for federal health programs [35]. The agency maintains a database of quality measures and indicators for many procedures. Currently, no endoscopy related quality measures have been endorsed by the NQF, primarily due to the absence of high-quality evidence of improved outcomes. Measurement of some of the outcomes of endoscopy is inherently challenging for many reasons. Some of the outcomes may not become apparent for a long period of time (development of malignancy after adenoma detection) or may be dependent on patient characteristics (comorbidity, adherence and socioeconomic factors) and disease severity which may not be amenable to risk adjustment. In the absence of direct clinical outcomes, surrogate markers (e. g. adenoma detection rate, withdrawal time, cecal intubation rate, and surveillance intervals) have sometimes been utilized to reflect the quality of care as process-based or structural measures. But in the absence of high-quality evidence, the use of such surrogate measures remains subject to bias [36].

The main limitation of the current systematic review is the significant variability in the reporting of these quality measures. Significant heterogeneity was observed in terms of both quality and quantity of metrics. Given the lack of uniformity in reporting outcomes i. e., some societies used standardized tools (such as GRADE and CEBM method), to attenuate this we provided a uniform perspective by using the pre-defined A, B, and C category system. Also, a few of the measures reported by societies were published 5 or more years ago. The authors acknowledge that age alone is not a fundamental limitation to guideline adherence so long as the evidence base is strong and the guideline panel has a method to provide up-to-date recommendations as new evidence emerges. Finally, guidelines and quality measures are also promoted by payors and regulatory groups, however, while these used to be publicly reported, there has been a shift toward societies increasingly recommending them and thus were not included to reflect clinician/

provider-available data. Despite these, the major strengths of our systematic review were the inclusion of numerous data points, comparing nine distinct society/organizations, and encompassing an aggregate total of 214 quality measures. Outcome measures are the foundation of credible structural and process-based measures, and our study underscores the significance of outcomes-based research in quality measures in endoscopy.

In summary, majority of quality measures (>70%) in endoscopy are based on low-quality evidence with significant heterogeneity observed in reporting from different societies/organizations. While there should be appreciation for the respective quality measures and as such we congratulate the numerous societies to make recommendations especially when data is scarce; our data calls for a need of high-quality studies examining patient-centered quality measures, the application of a standardized reporting method (such as GRADE), regular update of guidelines (based on newer evidence), as well as a strict adherence to protocol (COI disclosure, patient participation, etc.) for reporting quality measures in gastrointestinal endoscopy. We also suggest that societies focus primarily on the important metrics, namely—those that will be patient-centric and outcome driven—to best simplify the take-home measures that clinicians and endoscopy centers should strive to comply with.

Competing interests

The authors declare that they have no conflict of interest.

References

- [1] Weissman S, Goldowsky A, Mehta TI et al. Are quality metrics in inflammatory bowel disease rooted on substantial quality evidence? a systematic review *J Crohns Colitis* 2020; 16: 336–354
- [2] Gurudu SR, Ramirez FC. Quality metrics in endoscopy. *Gastroenterol Hepatol* 2013; 9: 228–233
- [3] Vadlamudi C, Brethauer S. Quality in endoscopy. *Surg Clin North Am* 2020; 100: 1021–1047
- [4] Minoli G, Meucci G, Prada A et al. Quality assurance and colonoscopy. *Endoscopy* 1999; 31: 522–527
- [5] Ball JE, Osbourne J, Jowett S et al. Quality improvement programme to achieve acceptable colonoscopy completion rates: prospective before and after study. *BMJ* 2004; 329: 665–667
- [6] Rex DK, Bond JH, Winawer S et al. Quality in the technical performance of colonoscopy and the continuous quality improvement process for colonoscopy: recommendations of the U.S. Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol* 2002; 97: 1296–1308
- [7] Rex DK, Petrini JL, Baron TH et al. Quality indicators for colonoscopy. *Am J Gastroenterol* 2006; 101: 873–885
- [8] Feuerstein JD, Akbari M, Gifford AE et al. Systematic review: the quality of the scientific evidence and conflicts of interest in international inflammatory bowel disease practice guidelines. *Aliment Pharmacol Ther* 2013; 37: 937–946
- [9] Sardar P, Giri J, Jaff MR et al. Strength of evidence underlying the american heart association/american college of cardiology guidelines on endovascular and surgical treatment of peripheral vascular disease. *Circ Cardiovasc Interv* 2019; 12: 1–8
- [10] Duarte-Garcia A, Zamore R, Wong JB. The evidence basis for the American College of Rheumatology Practice Guidelines. *JAMA Intern Med* 2018; 178: 146–148
- [11] Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ* 2002; 325: 961–964
- [12] Vanclooster A, Cassiman D, Steenberghe WV et al. The quality of hereditary haemochromatosis guidelines: A comparative analysis. *Clin Res Hepatol Gastroenterol* 2015; 39: 205–214
- [13] Brito JP, Domecq JP, Murad MH et al. The Endocrine Society Guidelines: when the confidence cart goes before the evidence horse. *J Clin Endocrinol Metab* 2013; 98: 3246–3252
- [14] Eubank BH, Mohtadi NG, Lafave MR. Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. *BMC Med Res Methodol* 2016; 16: 56
- [15] Moher D, Liberati A, Tetzlaff J et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009; doi:10.1371/journal.pmed.1000097
- [16] Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017; 358: j4008
- [17] Sharma P, Parasa S, Shaheen N. Developing quality metrics for upper endoscopy. *gastroenterology* 2020; 158: 9–13
- [18] Rex DK, Schoenfeld PS, Cohen J et al. Quality indicators for colonoscopy. *Am J Gastroenterol* 2015; 110: 72–90
- [19] Park WG, Shaheen NJ, Cohen J et al. Quality Indicators for EGD. *Am J Gastroenterol* 2015; 110: 60–71
- [20] Rizk MK, Sawhney MS, Cohen J et al. Quality indicators common to all GI endoscopic procedures. *Am J Gastroenterol* 2015; 110: 48–59
- [21] Calderwood AH, Day LW, Muthusamy VR et al. ASGE guideline for infection control during GI endoscopy. *Gastrointest Endosc* 2018; 87: 1167–1179
- [22] Day LW, Cohen J, Greenwald D et al. Quality indicators for gastrointestinal endoscopy units. *VideoGIE* 2017; 26: 119–140
- [23] Park WG, Shaheen NJ, Cohen J et al. Quality indicators for EGD. *Gastrointest Endosc* 2015; 81: 17–30
- [24] Beg S, Ragnunath K, Wyman A et al. Quality standards in upper gastrointestinal endoscopy: a position statement of the British Society of Gastroenterology (BSG) and Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland (AUGIS). *Gut* 2017; 66: 1886–1899
- [25] Armstrong D, Barkun A, Bridges R et al. Canadian Association of Gastroenterology consensus guidelines on safety and quality indicators in endoscopy. *Can J Gastroenterol* 2012; 26: 17–31
- [26] Kaminski MF, Thomas-Gibson S, Bugajski M et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscopy* 2017; 49: 378–397
- [27] Valori R, Cortas G, de Lange T et al. Performance measures for endoscopy services: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscopy* 2018; 50: 1186–1204
- [28] Atkin WS, Valori R, Kuipers EJ et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition-Colonoscopy surveillance following adenoma removal. *Endoscopy* 2012; 44: SE151–SE163
- [29] Fletcher RH, Nadel MR, Allen JI et al. The quality of colonoscopy services—responsibilities of referring clinicians: a consensus statement of the Quality Assurance Task Group, National Colorectal Cancer Roundtable. *J Gen Intern Med* 2010; 25: 1230–1234
- [30] Grupo de Trabajo de “Indicadores de calidad en endoscopia” de la Sociedad Española de Patología Digestiva (SEPD). Quality indicators

- in colonoscopy. The colonoscopy procedure. *Rev Esp Enferm Dig* 2018; 110: 316–326
- [31] Jover R, Herráiz M, Alarcón O et al. Clinical practice guidelines: quality of colonoscopy in colorectal cancer screening. *Endoscopy* 2012; 44: 444–451
- [32] Chassin MR, Galvin RW. The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality. *JAMA* 1998; 280: 1000–1005
- [33] Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington (DC): National Academies Press (US); 2001
- [34] Guyatt GH, Oxman AD, Vist GE et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336: 924–926
- [35] Institute of Medicine (US) Committee on Conflict of Interest in Medical Research, Education, and Practice. Lo B, Field MJ. *Conflict of Interest in Medical Research, Education, and Practice. Conflicts of Interest and Development of Clinical Practice Guidelines*. Washington (DC): National Academies Press (US); 2009; 7: <https://www.ncbi.nlm.nih.gov/books/NBK22928/>
- [36] Panzer RJ, Gitomer RS, Greene WH et al. increasing demands for quality measurement. *JAMA* 2013; 310: 1971–1980