



# Quality Improvement in Gastroenterology: A Systematic Review of Practical Interventions for Clinicians

Courtney Reynolds<sup>1</sup> · Eric Esrailian<sup>1,2</sup> · Daniel Hommes<sup>1</sup>

Received: 26 February 2018 / Accepted: 5 July 2018 / Published online: 16 July 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

**Background** Quality improvement (QI) identifies practical methods to improve patient care; however, it is not always widely known which QI methods are successful. We sought to create a primer of QI in gastroenterology for the practicing clinician.

**Methods** We performed a systematic review of QI literature in gastroenterology. We included search terms for inflammatory bowel disease, irritable bowel syndrome, celiac disease, gastroesophageal reflux disease, pancreatitis, liver disease, colorectal cancer screening, endoscopy, and gastrointestinal bleeding. We used general search terms for QI as well as specific terms to capture established quality metrics for each GI disease area.

**Results** We found 33 studies that met our definitions for QI. There were 17 studies of endoscopy including screening colonoscopy, six on liver disease, four on IBD, two on GERD, three on GI bleeding, and one on celiac disease. Education was the most common intervention, although most successful studies combined education with another intervention. Other effective interventions included retraining sessions to reach ADR goals in colonoscopy, nursing protocols to increase HCC screening, and EMR decision support tools to prompt reassessment of PPI therapy. Many studies showed improved compliance to metrics, but few were able to show differences in length of stay, readmissions, or mortality.

**Conclusions** Our review of quality improvement literature in gastroenterology revealed common themes of successful programs: Education was frequently used but often insufficient, the EMR may be underutilized in guiding decision making, and patient-reported outcomes were infrequently assessed. Further research may be needed to compare QI strategies directly.

**Keywords** Quality improvement · Health services · Gastroenterology · Delivery of health care

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10620-018-5198-x>) contains supplementary material, which is available to authorized users.

---

✉ Courtney Reynolds  
careynolds@mednet.ucla.edu

Eric Esrailian  
eesrailian@mednet.ucla.edu

Daniel Hommes  
dhommes@mednet.ucla.edu

<sup>1</sup> Division of Digestive Diseases, Department of Medicine, University of California Los Angeles, 10945 Le Conte Ave, Suite 2338, Los Angeles, CA 90095, USA

<sup>2</sup> Division of Digestive Diseases, Department of Medicine, University of California Los Angeles, 200 Medical Plaza Driveway, Los Angeles, CA 90024, USA

## Introduction

The rising cost of health care has been a driving force behind the nascent field of quality improvement. In the USA, health-care spending represents 17% of the gross domestic product (GDP) and is expected to increase by as much as 5% each year in the coming decade [1, 2]. This trend is in sharp contrast with the majority of the world's developed nations—including Canada, France, Sweden, Japan, the UK and Australia—which spend a stable 10% or less of their GDP on health care [1]. Notably, higher healthcare spending in the USA has not resulted in better health outcomes, which is reflected in high maternal mortality, high infant mortality, and decreasing life expectancy [3–5].

The Affordable Care Act of 2010 contained several provisions for improving quality in health care, but the most visible effort has been the Center for Medicare and Medicaid Services hospital readmissions reduction program [6]. The result has been an explosion of quality improvement efforts

primarily focused on reducing readmissions for heart failure, myocardial infarction, and pneumonia. In particular, a brief review of the literature on quality improvement shows that the majority of published studies focus on heart failure and the transition from acute care to outpatient care, with primary outcomes of readmission rate and cost [7–11]. Readmissions are often treated as a substitute metric for quality.

As quality improvement projects have become more popular and numerous, there is a need to examine which strategies have proven successful in order to guide future efforts and facilitate implementation at clinician practices and departments. Our field of gastroenterology in particular has witnessed a recent increase in QI-related publications. Here, we aimed to compile a practical guide to successful QI within gastroenterology, focusing upon established quality metrics for a range of acute and chronic diseases. We conducted a systematic review of the literature.

## Methods

### Defining Quality Improvement in Gastroenterology

To frame our literature search, we used a commonly accepted definition of quality improvement (QI): a systematic approach to the analysis of practice performance and efforts to improve performance. In the context of health care, we viewed QI as a means of improving both process and outcomes. Accordingly, we did not include studies that simply assessed how well we are currently meeting quality standards, but restricted our results to studies with an active intervention [12]. To obtain quality metrics for specific GI diseases, we reviewed current disease guidelines (ACG, AGA, AASLD), Choosing Wisely, and the Physician Quality Reporting System (PQRS) [13, 14].

### Systematic Review

We performed a systematic search of the QI literature within gastroenterology. Our PubMed/Medline search included terms for specific diseases within gastroenterology: inflammatory bowel disease (IBD), irritable bowel syndrome (IBS), celiac disease, gastroesophageal reflux disease (GERD), chronic and acute pancreatitis, chronic liver disease and cirrhosis, colorectal cancer screening, endoscopy, and gastrointestinal bleeding. Our terms for quality improvement included: health services research, healthcare delivery, healthcare improvement, clinical practice, quality of care, quality indicators, quality improvement, quality metrics, and reducing readmissions (please see Supplement 1 for exact search terms). We restricted our search to studies performed on adults and in humans. We included only experimental and quasi-experimental studies and excluded studies of surgical

treatment or surgical outcomes. While we restricted our search to studies written in English, we did include studies done in other countries. We conducted a search of the grey literature, including OpenGrey, Health Systems Evidence, and the New York Academy of Grey Medicine Literature.

## Results

### Systematic Review

Our initial search of PubMed yielded 5345 results (see Supplement 1 for search terms). After title, abstract, and full text review cycles, we refined this list to 28 studies. We added five more studies found by hand search, from either searching the references of already included studies, or Google. Our search of the grey literature did not yield any additional titles that met our criteria.

Descriptive characteristics for each of the 33 studies are summarized in Table 1. The majority of studies were conducted in the USA, after passage of the Affordable Care Act. There were 17 studies on endoscopy, six on chronic liver disease including cirrhosis, four on IBD, three on GI bleeding, two on GERD, and one on celiac disease. We did not find any relevant studies of IBS or pancreatitis.

### Celiac Disease

There are no formally established quality metrics for celiac disease. To guide our search, we reviewed the ACG guidelines for diagnosis and management of celiac disease and searched for QI papers that would address appropriate serologic testing, taking adequate biopsies during EGD, or referring to a nutritionist to gluten free diet counseling. The one study we found used a pre-/posttest design to encourage providers to order tissue transglutaminase antibody (TTG) alone for serologic diagnosis, as compared to a “celiac disease panel” that included TTG as well as anti-gliadin antibody (GA IgG and IgA) and endomysial antibody (EMA IgG/IgA) [15]. Guidelines recommend TTG testing alone for initial testing for celiac disease. Based upon discussion with gastroenterologists, the study initiated the following interventions: eliminating the celiac panel order and educating physicians of expected first-line and second-line tests through a laboratory memo. After this intervention, repeat audit of laboratory testing showed that the rate of TTG alone increased and rates of ordering GA or EMA decreased significantly. While this study used education as a means to reach out to ordering physicians, this intervention was also paired with changing the EMR to encourage following the guidelines. As in the majority of studies in this review, education is a common component of QI but is often insufficient on its own.

**Table 1** Description of quality improvement studies in gastroenterology

Article	Quality metric	Study design	Sample size and population	Intervention	Outcome
<i>Celiac disease</i>					
Huang 2012	TTG testing	Pretest/posttest	Physician orders to screen celiac disease; all ages	Direct education to physicians on ordering guidelines	Positive; increased rate of ordering TTG alone versus celiac panel
<i>Colorectal cancer screening and colonoscopy</i>					
Coe 2013	ADR	RCT	15 academic gastroenterologists	EQUIP: randomized to 2 h retraining sessions plus individual feedback	Positive; increased ADR
Ussui 2015	ADR	Pretest/posttest	15 academic gastroenterologists	EQUIP2: follow-up study of Coe to see whether higher ADR sustained over time	Positive; sustained increase in ADR at 5 months
Wallace 2017	ADR	RCT	Nine practices; 22,316 colonoscopies	EQUIP3: multicenter version of EQUIP retraining and feedback	Equivocal; increased ADR at both control and intervention sites
Inra 2017	ADR, CIR, WT	Pretest/posttest	Three practices; 28 endoscopies	Individual performance scorecard	Negative; no change in endoscopy metrics 12 weeks after scorecard
Imperiali 2006	ADR, CIR	Interrupted time series	Single endoscopy unit; over 10,000 endoscopies	Individual and group feedback, supervision by more experienced endoscopists	Partially positive; ADR unchanged, completion rate improved
Kahi 2013	ADR, CIR	Pretest/posttest	Single VA medical center; 928 endoscopies	Feedback through quarterly report card	Partially positive; ADR unchanged, completion rate improved
Shaukat 2009	ADR	Interrupted time series	Five ambulatory endoscopy centers; 47,253 endoscopies	Financial penalty for WT, retraining for poor performers, individual and group feedback	Negative; wide variation in ADR by individual, no change over 3 years
Ross 2015	ADR, SSADR	Pretest/posttest	Single tertiary cancer center; 2833 endoscopies	Financial incentive for ADR plus individual and group feedback	Equivocal; nonsignificant increase in both ADR and SSADR
Barclay 2008	WT, ADR	Pretest/posttest	Single community GI practice; 12 providers	Eight-minute timer set to encourage minimum withdrawal time	Positive; increased average withdrawal time and ADR
Rajasekhar 2015	Colonoscopy technique bundle	Pretest/posttest	12 community endoscopy units; 17,508 endoscopies	Education and training for adherence to bundle: WT, supine positioning, rectal retroflexion, use hyoscine butylbromide	Positive; increase in ADR and hyoscine use, but unclear results for WT or positioning
Grassini 2008	Inappropriate colonoscopy referrals	Pretest/posttest	Single endoscopy unit; 1017 endoscopies	Education of family physicians in an open-access system through a CME course	Positive; decreased rate of inappropriate referrals for colonoscopy
Mai 1991	Documentation for EGD and colonoscopy	Pretest/posttest	Single VA medical center; 1408 endoscopies	Monthly individual feedback, group discussion	Positive; increased rate of complete documentation per ASGE guidelines
Abdul-Baki 2016	ADR	Pretest/posttest	Single community GI practice; 20,040 endoscopies	Public reporting of “colonoscopy quality index” based on ADR, adverse events, documentation	Positive; improved ADR

Table 1 (continued)

Article	Quality metric	Study design	Sample size and population	Intervention	Outcome
Hayat 2016	Bowel prep quality, ADR	Prospective cohort	Cleveland Clinic; 2530 patients	Patient video on bowel prep (30 min)	Partially positive; improved bowel prep but ADR rates the same
Hsueh 2014	Bowel prep quality	Non-randomized controlled trial	Single community hospital; 218 patients	Patient video on bowel prep (8 min)	Positive; improved bowel prep
Clayton 2016	Bowel prep quality	Pretest/posttest	Academic liver transplant clinic; 121 patients	Patient video on bowel prep (6 min)	Negative; no change in bowel prep quality
Park 2016	Bowel prep quality; WT, ADR	RCT	Single academic center; 502 patients	Patient video on bowel prep (6 min)	Partially positive; improved bowel prep but ADR rates the same
<i>Gastroesophageal reflux disease (GERD)</i>					
Walsh 2016	De-prescribing of inappropriate PPI	Pretest/posttest	Family medicine practice	EMR message to reassess PPI necessity if used over 8 weeks	Positive; > 90% chronic PPI users reassessed, about 25% de-prescribed
Player 2010	Appropriate diagnosis and treatment initiation	RCT	27 primary care offices; 119 providers	EMR message to consider GERD diagnosis or treatment based on ICD-9 codes for GERD, asthma, chronic cough	Positive; higher rate of GERD diagnosis and initiation of treatment in intervention group
<i>Gastrointestinal bleeding</i>					
Johnson	Readmissions	Pretest/posttest	Single tertiary care center; 99 patients	EMR order set to encourage optimal care variceal bleeds; somatostatin analogue, PPI, antibiotics	Positive; reduced readmissions with increased compliance to optimal care
Loftus	Readmissions, LOS	Pretest/posttest	Single hospital; 237 patients	Mandatory meeting between GI, IR and surgery for severe GI bleeding	Positive; shorter LOS, reduced readmissions, earlier intervention and fewer blood transfusion
Pfau	Cost, LOS, time to endoscopy, use of IV versus PO acid blockers	Pretest/posttest	Single tertiary care center; 421 patients	Standard care pathway for upper non-variceal GI bleeding, disseminated through e-mails and lectures	Partially positive; reduce use of IV acid blocking agents, but no changes in LOS, cost or time to endoscopy
<i>Inflammatory bowel disease (IBD)</i>					
Walsh 2013	Vaccination rates	Pretest/posttest	30 gastroenterologists from eight different centers	Education through a handout on current vaccine guidelines	Positive; higher rates of vaccination for flu, hepatitis, pneumonia
Sapir 2016	PQRS measures for IBD	Pretest/posttest	40 community-based gastroenterologists	Education through group chart review	Positive; improved vaccination rates, testing for TB and hepatitis B, among others
Greene 2015	PQRS measures for UC	Pretest/posttest	30 community-based gastroenterologists	Education and individual feedback through chart audits	Positive; improved self-reported understanding of measures
Greene 2015	PQRS measures for Crohn's disease	Pretest/posttest	20 community-based gastroenterologists	Education and individual feedback through chart audits	Partially positive; improved adherence to four of ten quality measures

**Table 1** (continued)

Article	Quality metric	Study design	Sample size and population	Intervention	Outcome
<i>Chronic liver disease</i>					
Loy 2016	Preventive care for cirrhosis	Interrupted time series	Single academic GI clinic; 2311 patients	Education plus individual feedback	Positive; improved rates of hepatitis vaccination, screening for varices and HCC
Kennedy 2013	HCC screening in chronic viral hepatitis	Interrupted time series	Single tertiary care center; 114 patients	Education to both patients and providers; dedicated nurse to identify relevant patients and contact them	Positive; increased rate of biannual HCC screening
Ghaoui 2015	LOS, readmissions, mortality	Pretest/posttest	Single tertiary care center; 303 patients	Mandatory GI consultation for admitted patients with decompensated cirrhosis	Partially positive; improved composite QI score based on early paracentesis, optimal care of GI bleeding, documented transplant evaluation; no change in LOS, readmissions, mortality
Desai 2014	LOS, readmission, mortality	Pretest/posttest	Single tertiary care center; 56 patients	Co-management of admitted liver disease patients by hepatology	Partially positive; no difference in LOS, mortality, readmissions, but improved adherence to optimal care such as early paracentesis or prophylaxis for SBP
Aberra 2013	HCC screening in cirrhosis	Pretest/posttest	Single tertiary care center; 362 patients	Nursing-based protocol for automatic reminders	Positive; increased rates of timely HCC screening
Tapper 2017	Readmissions	Pretest/posttest	Single tertiary care center; 824 patients	EMR checklist to guide inpatient management of cirrhosis	Positive; lower readmissions, and higher rates of lactulose and rifaximin use for HE, chemical VTE prophylaxis, SBP treatment and beta-blocker use

*TTG* tissue transglutaminase, *ADR* adenoma detection rate, *CIR* cecal intubation rate, *ASGE* American Society for Gastrointestinal Endoscopy, *SSADR* sessile serrated adenoma detection rate, *GERD* gastroesophageal reflux disease, *PPI* proton pump inhibitor, *EMR* electronic medical record, *GI* gastrointestinal, *LOS* length of stay, *IR* interventional radiology, *IBD* inflammatory bowel disease, *PQRS* physician quality reporting system, *TB* tuberculosis, *UC* ulcerative colitis, *HCC* hepatocellular carcinoma, *QI* quality improvement, *SBP* spontaneous bacterial prophylaxis, *HE* hepatic encephalopathy, *VTE* venous thromboembolism

## Chronic Liver Disease and Cirrhosis

Studies of liver disease were evenly divided among the inpatient and outpatient setting. Among inpatient studies, Tapper developed an electronic checklist for use in the hepatology ward, to increase compliance to guidelines for management of cirrhosis and its complications [16]. Guidelines covered inpatient acute issues such as chemical prophylaxis for venous thromboembolism and antibiotic prophylaxis for spontaneous bacterial peritonitis, but also included initiation of long-term therapy to continue post-discharge such as beta-blockers for varices and lactulose and rifaximin for hepatic encephalopathy. Ultimately, the electronic checklist resulted in more patients receiving secondary prophylaxis for spontaneous bacterial peritonitis and appropriate dosing of lactulose, and its use was associated with significantly lower 30-day readmissions and in readmissions specifically for hepatic encephalopathy. Of note, all patients with hepatic encephalopathy in the checklist study were routinely placed on rifaximin. The Ghaoui and Desai studies each created protocols for mandatory GI involvement for admitted patients with liver disease [17, 18]. Ghaoui restricted this initiative to patients with decompensated cirrhosis, while Desai went further by dictating hospitalist and hepatologist co-management for all liver disease patients. Despite these differences, both studies found significantly increased compliance to quality measures for liver disease, including early paracentesis, prophylactic antibiotics for SBP or GI bleeding, early endoscopy for GI bleeding, and appropriate use of FFP (Table 2).

In the outpatient setting, the Kennedy and Aberra studies both targeted screening for hepatocellular carcinoma (HCC); the patient population in Kennedy was chronic viral hepatitis, versus cirrhosis for Aberra [19, 20]. Both studies utilized nursing-based protocols to improve screening rates; in these protocols, patients were entered into a separate database and contacted for HCC screening independent of their clinic visits. Further, nurses were able to order ultrasounds directly and to schedule reminder calls. Both programs, which were remarkably similar, showed significantly increased screening rates post-intervention. Similarly, the Loy study sought to improve HCC screening rates in patients with cirrhosis but also included hepatitis A and B vaccination and screening for esophageal varices [21]. In contrast to Kennedy and Aberra, the goal of the Loy study was to increase provider rather than patient compliance with cirrhosis quality measures, using individual feedback. Provider feedback was given at baseline following chart review, and this process was repeated intermittently over 3 years. Using this continuous QI approach, Loy was able to demonstrate a significant increase in compliance with vaccination and HCC screening starting at 2 months that was sustained through 3 years; screening for esophageal varices was not affected.

Among the liver disease studies, there were no RCTs. Study designs were split between interrupted time series (Loy, Kennedy) and pre–posttest (all others), the most common designs for QI. All studies were performed at a single center, and sample size varied from 56 patients to over 2000. However, there was no clear relationship between sample size and how successful the study was, likely due

**Table 2** Selection of useful strategies for QI

QI metric or disease area	Strategies	Comments and pitfalls
Adenoma detection rate	Retraining or supervision Public reporting	Target poorly performing endoscopists Unclear whether public reporting would be interpreted correctly by patients Unclear whether financial incentives effective
Appropriate endoscopy referrals	Educating referring physicians: offer incentive such as CME credits	Time-intensive, only applicable to open-access system
Appropriate use of PPIs	EMR tools to prompt initiation and reassessment of PPI use	Alert fatigue among physicians; correctly identifying patients based on ICD-9 coding
Bowel prep for colonoscopy	Patient education videos	Unclear what length is optimal; better to be able to view at home versus office
GI bleeding	Standardized care pathway Predetermined interdisciplinary management	Simplicity is best for care pathways Interdisciplinary management approach likely to vary by institution
Hepatocellular carcinoma screening	Dedicated nurse and protocol to identify patients independent of physician visits	Cost of dedicated staff
Inflammatory bowel disease	Direct physician education and chart review	The educational interventions here also had an underlying financial incentive
Inpatient liver disease care	Mandatory consultation or co-management by GI	Increased workload for physician on service

CRC colorectal cancer, PPI proton pump inhibitors, EMR electronic medical record

to heterogeneity in the study populations (decompensated, inpatients, vs. stable outpatients) and in interventions. Impressively, each study showed increased compliance to the disease-specific quality measure of choice, but none were able to directly demonstrate effects on mortality, length of stay, or readmissions. The most likely reason for this is unaccounted confounding; the role of sample size is unclear. As in most QI studies, there were no sample size calculations, and it was unclear what factors were considered for sample size.

## Endoscopy

The 17 endoscopy studies can be grouped into those focusing on provider metrics—including adenoma detection rate (ADR), withdrawal time (WT), documentation, and referrals—and those focusing on patient metrics, such as bowel prep quality. The most frequent study design, used by ten papers, was pretest/posttest, followed by three randomized controlled trials, two interrupted times series, one non-randomized controlled trial, and one prospective cohort study. Units of study included individual providers, practices/clinics, and endoscopies; sample sizes ranged from 10 to 15 providers at a single clinic to over 10,000 endoscopies. Despite the variations in design, intervention, and sample size, only three of the 17 studies were negative; both negative studies of ADR looked at individual changes in performance rather than aggregate performance (Inra, Shaukat), while the negative study of bowel prep solely recruited patients with decompensated cirrhosis (Clayton) [22–24].

Among the 12 studies on provider metrics, the most common metric was ADR and the most common intervention was individual and group feedback. However, only three studies—Inra, Mai, and Kahi—had education as their sole intervention [22, 25, 26]. Neither Inra nor Kahi showed differences in ADR; Mai resulted in improved documentation for EGD and colonoscopy. The studies that also incorporated retraining, financial incentives, or other active interventions were more successful: Abdul-Baki et al. [27] utilized public reporting of endoscopists' quality metrics, including rate of complications, appropriate documentation, correct recommendation for follow-up, withdrawal time, and adenoma detection rate. The study then followed participating gastroenterologists for 4 years after the public report and demonstrated a significant increase in average ADR, albeit without a concurrent control group. In Ross, physicians at a single center received their individual versus practice average ADR and sessile serrated adenoma detection rates (SSA) and provided a financial incentive based upon productivity, ADR and SSA [28]. They were able to show a nonsignificant increase in both ADR and SSA. The EQUIP series of studies deployed a program of video-based education, individual feedback, and direct retraining sessions

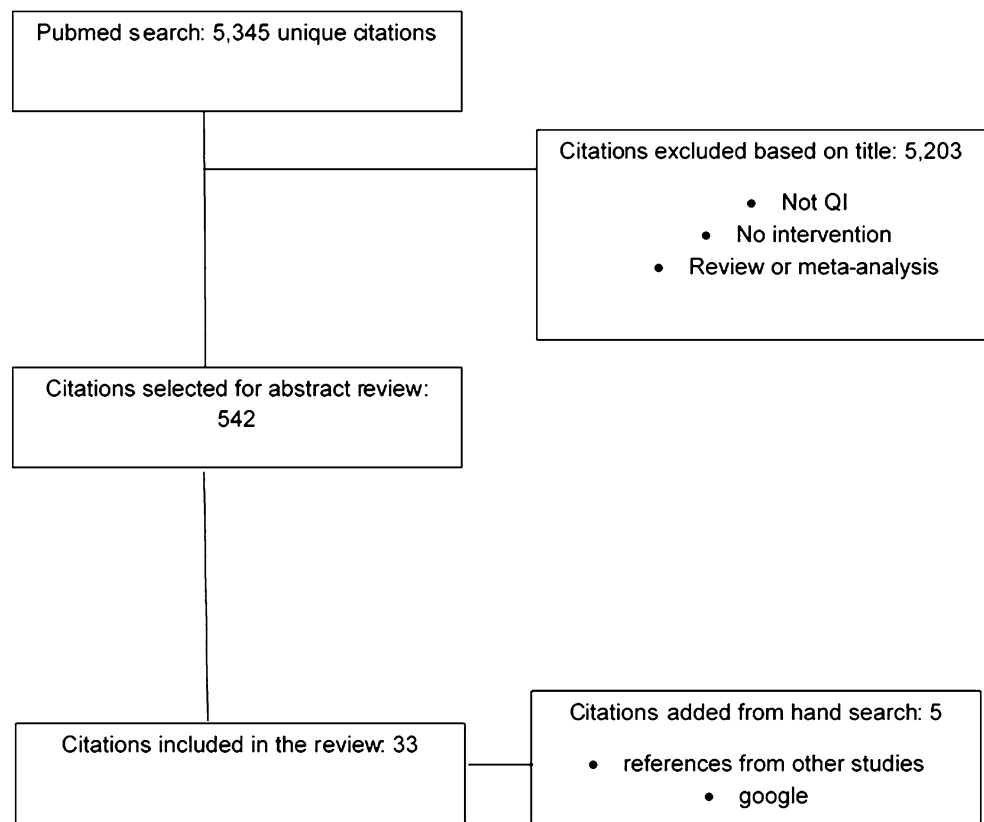
in endoscopy techniques [29–31]. The initial single-center study showed a significant increase in ADR (Coe), which was sustained over several months in the follow-up study by Ussui. However, the subsequent multicenter RCT done by Wallace was equivocal, showing increased ADR in both control and intervention sites. Like EQUIP, the Shaukat study also used direct retraining sessions but additionally included a financial penalty for not reaching minimum ADR and withdrawal time metrics; nonetheless, no significant difference was found from this intervention [23]. Of note, Shaukat looked for differences in individual-level quality metrics, rather than an aggregate before and after ADR across providers. The results of the Shaukat study raise an interesting question: Does imposing a financial penalty, instead of a financial reward, hurt performance by harming physician morale?

In Rajasekhar, the study group trained “local leaders” on their colonoscopy technique bundle, which included supine positioning, rectal retroflexion, and reaching minimum WT; these leaders were then tasked to hold training sessions for clinic staff and send periodic reminders about adherence [32]. The study was positive but its results are difficult to interpret however, as compliance with a single one of the bundle's measures was used as a surrogate for overall compliance with all measures. Still, this “trickle down” training approach may be a cost- and time-efficient way to disseminate education and retraining. The Imperiali et al. [33] study implemented a “continuous quality improvement” approach in which they serially measured colonoscopy completion rates and ADR in 6-month audit cycles and identified several points to intervene. Ultimately, they changed sedation practices to increase patient comfort, gave lower performing physicians access to training sessions, and continually gave individual feedback. At study end, all endoscopists had reached over 90% completion rate and 20% ADR (Fig. 1).

As their sole intervention, Barclay placed a timer set to 8 min in the endoscopy suite to ensure the minimum withdrawal time; in addition to sounding at 8 min, there were intermediate alarms set to help endoscopists pace their procedure (the timer could be paused while taking biopsies) [34]. The timer system led to significantly increased withdrawal time and ADR. However, the long-term effectiveness of this approach is unclear, as it would be impractical, or at least irritating, to have a timer present during procedures indefinitely.

Four studies (Clayton, Hayat, Park, and Hsueh) focused upon patient education to improve bowel prep during outpatient screening colonoscopy [24, 35–37]. Each study created a patient education video, ranging from 6 to 30 min, explaining the importance of bowel prep in successfully completing colonoscopy. Despite this similarity, one study did not show any change in bowel prep while the others did. In the unsuccessful study, Clayton, there

Fig. 1 Study selection



was a relatively small sample size but more importantly, the video was played in the office during the course of a visit for liver transplantation evaluation rather than at home at the patient's leisure. These patients were higher acuity in general and may have had more difficulty with bowel prep than the average patient; patients also had some degree of ascites, as an additional challenge. Two of the positive studies, Hayat and Park, also examined whether improved bowel prep leads to increased ADR, but both studies were negative on this front. The length of the video varied significantly but did not appear to affect the success of the study.

The final endoscopy study, Grassini, targeted inappropriate referrals for colonoscopy [38]. Family physicians in an open-access endoscopy system were educated on criteria for appropriate colonoscopy referrals via a full-day continuing medical education (CME) course, along with a follow-up letter repeating the list of approved criteria. Gastroenterologists rated the appropriateness of referrals they received before and after this intervention and found a significant improvement. While this was an education-based intervention, the Grassini study wisely offered providers CME, which likely led to increased attendance and engagement.

### Gastroesophageal Reflux Disease and Proton Pump Inhibitor Use

Both studies in this section occurred in the clinic setting and included measures for appropriate use of proton pump inhibitor (PPI) therapy. The Player study was an RCT to test an EMR-based decision support tool encouraging physicians to initiate PPI therapy if GERD had been diagnosed previously and to consider the diagnosis of GERD if atypical symptoms such as chronic cough and asthma were present (based on ICD-9 coding) [39]. Implementation of the tool increased the rate of new GERD diagnoses and the odds of initiating treatment for those with atypical symptoms. At the other end of the care pathway, the Walsh study promoted reassessment of patients on chronic PPI therapy, with the goal of de-escalating or de-prescribing if appropriate [40]. Their study also used the EMR, to first deliver a message to reassess therapy in patients on PPI for greater than 8 weeks and then to present a decision support tool to guide either continuation, lower dosing, or complete discontinuation of PPI. The study used a pretest/posttest design; subsequent chart review showed the need for PPI was reassessed in over 90% of these patients, and about 25% were de-prescribed. Both studies showed that the EMR can be successfully leveraged



to improve adherence to guidelines, although the cumulative effect of multiple EMR-based clinical decision support tools may lead to “click fatigue” among providers.

## Gastrointestinal Bleeding

Although several studies included GI bleeding among their outcomes, three studies focused exclusively on GI bleeding. The Loftus study initiated a protocol for a mandatory conference call between consultants from GI, interventional radiology, and surgery in the case of severe GI bleeding, defined by: large-volume bleeding, hemodynamic instability, four or more units red blood cells transfused within 24 h or 8 units total, history of recurrent bleed, re-bleeding after endoscopy or no clear source of continued bleeding on endoscopy, patient is Jehovah’s witness refusing transfusion, or at GI service’s discretion [41]. This multi-disciplinary approach led to shorter time to procedures (whether endoscopic, surgical, or radiographic) and fewer transfusions and was associated with shorter LOS and lower readmissions for recurrent GI bleeding.

The Johnson study focused on GI bleeding within the subset of suspect variceal bleed in cirrhosis [42]. Initially, they attended departmental conferences to educate hospitalists and intensivists on “optimal care,” defined by use of a proton pump inhibitor, somatostatin analogue, and prophylactic antibiotics. An electronic admission order set was then implemented which led to significantly more patients receiving optimal care and to lower 30-day readmissions. LOS was unaffected, and the study was not sufficiently powered to test the effect on mortality.

Finally, Pfau debuted a standard care pathway for non-variceal upper GI bleeding, including type and timing of laboratories, when to obtain X-ray imaging, use of IV versus PO acid blocking medications, and indications for ICU admission or for discharge [43]. The standardized pathway was developed by an interdisciplinary hospital panel and disseminated through e-mails and lectures. Their outcomes were time to endoscopy, length of stay, inappropriate use of X-rays, and inappropriate use of IV H2 blockers or PPIs. In a pretest/posttest study design, they were able to show reduced rates of inappropriate IV H2 blockers or PPIs, but no difference in other outcomes including LOS or time to endoscopy.

All three GI bleeding studies used a pretest/posttest design and occurred at a single hospital center. Johnson and Loftus were able to show both compliance with their interventions and resulting improvements in readmission rate and LOS. Pfau, while having the largest sample size, was not able to show an effect on these quality metrics; one reason may be that their standardized care pathway was simply too long, discouraging adoption. Their pathway had seven categories (such as physical examination, criteria for ICU admission) with 2–6 items in each category. In addition, several

items on the pathway were appropriate and may reduce cost but probably do not affect clinical outcomes, including the use of PO instead of IV acid blockers. It is possible that the most important interventions were lost in the excessively comprehensive set of recommendations. However, overall these studies showed that checklists and other care guidelines can improve outcomes for critically ill patients.

## Inflammatory Bowel Disease

All four studies of IBD had pretest/posttest design and studied ways to improve adherence to IBD guidelines among a small group of community-based physicians. Three of the four IBD studies originated from the same research group and focused on increasing understanding and adherence for Physician Quality Report System (PQRS) and National Quality Strategy (NQS) measures. PQRS measures for IBD include minimizing corticosteroid use, performing appropriate vaccinations, and testing for hepatitis B and tuberculosis prior to initiating anti-TNF therapy [14]. The six NQS priorities are broader and non-disease specific, encompassing patient safety, patient-centered care, care coordination, preventative care, lifestyle changes, and cost effectiveness [44].

The Sapir study provided education to gastroenterologists through group chart review (which counted as continuing medical education) and included a control group [45]. Subsequent chart audit showed significantly increased adherence for the intervention group across several measures including influenza and pneumonia vaccination, testing for hepatitis B and tuberculosis, and assessment of treatment side effects.

The two Greene papers studied the same PQRS measures for IBD, but separately in the UC versus Crohn’s populations; each study used education as well as chart auditing with individual physician feedback [46, 47]. The Greene UC study showed increased understanding of the measures and intent to apply them, but did not actually perform repeat chart audit or other follow-up to measure post-intervention compliance. However, their later Crohn’s study did include a post-intervention repeat chart audit. While there was no overall difference in adherence before and after intervention, a secondary analysis showed significant improvement in four of ten measures among initially “low performing” physicians: assessment of IBD type, assessment of IBD activity, testing for bone loss from corticosteroid use, and tobacco cessation counseling.

Finally, the Walsh study developed a proforma covering vaccination and screening for several infections in IBD: HPV, HIV, hepatitis C, varicella, influenza, pneumococcus and tuberculosis [48]. Gastroenterologists were surveyed for their baseline awareness of the relevant guidelines and their current practice; charts were also audited to determine baseline and post-intervention screening and vaccination rates. However, these post-intervention rates were

not communicated to physicians. Following this education, physicians self-reported greater compliance to guidelines, while chart review showed increased orders and referrals placed for these vaccinations and testing. Overall, the studies were positive and showed higher adherence to guidelines after physician education. This is somewhat surprising, since many QI studies using education as the sole intervention have not been successful. It is possible that physicians are more likely to be interested in adhering to PQRS measures since there is a financial incentive for doing so.

## Discussion

Our review of quality improvement literature in gastroenterology revealed several useful takeaway points to guide departments and clinicians in implementation. First, education was the most widespread and simple intervention used in these studies; however, its success was highly variable. This finding is likely the result of the heterogeneity contained within education as an intervention and the fact that many studies paired education with another intervention. For example, education in some of the colonoscopy studies could be as simple as giving physicians a “report card” on their performance meeting adenoma detection rates; this approach was generally unsuccessful unless coupled with another tactic such as retraining sessions for low-performing endoscopists. In contrast, the Grassini study also used education (successfully) to reduce the rate of inappropriate referrals for colonoscopy in an open-access system. However, here education referred to a full-day information session for family medicine clinicians that granted them several CME credits. The (largely) successful studies to improve compliance with PQRS measures in IBD utilized education, but improved compliance with PQRS also has a potential financial benefit for physicians treating Medicare patients. While education is a viable, low-cost strategy in quality improvement, the likelihood of success is directly related to the quality of the education and whether it is paired with any other interventions. It would be helpful to use education as an “active control” in future QI studies in order to quantify what contribution education makes to the success of a QI project when paired with other interventions.

Other successful QI programs used the electronic medical record to increase physician compliance with high-quality and guideline-based care. Tapper used an EMR checklist to improve inpatient cirrhosis care, Johnson implemented an EMR order set to outline “optimal care” for variceal bleeding, and both GERD studies (Walsh and Player) used decision support tools to encourage appropriate initiation and discontinuation of proton pump inhibitors. All four of these studies were quite effective and had the advantage of not requiring any additional staff time or resources to operate

continuously once implemented. On the other hand, widespread application of EMR messages, mandatory forms, and decision support tools has the potential if used too frequently to inspire user fatigue. In a given system, the total number of EMR tools used must be carefully considered to avoid habitual “alert override,” a rising problem in which clinicians become inundated with alerts and began to ignore them indiscriminately, with a resulting decreased in compliance and quality [49]. Another consideration for EMR interventions is the level of complexity. The Pfau GI bleeding study was less successful than Tapper or Johnson, and one potential reason is the length and scope of their standardized care pathway for upper GI bleeding. The EMR itself has added complexity to providers’ tasks, and any interventions utilizing the EMR should be as streamlined as possible to increase the likelihood of adherence.

To improve patient compliance, dedicated staff and protocols appear to have merit. The Kennedy and Aberra studies of HCC screening in cirrhosis were remarkably similar, and both utilized nurse-driven protocols to identify patients in need of screening and remind patients by phone or messaging. Both studies were effective and circumvented the common problem of patient no-shows or lack of follow-up: Patients were contacted independently of office visits and nurses were able to order ultrasounds. While this approach is labor-intensive, employing dedicated staff may well be cost-effective when used to prevent highly morbid outcomes such as diagnosis of advanced HCC.

Another “low tech” strategy for QI was to mandate GI and other sub-specialty involvement in the management of select patient populations. Loftus required a multi-disciplinary meeting among GI, IR, and surgery for severe GI bleeding, and Ghaoui implemented GI consultation for all inpatients with decompensated cirrhosis. Desai went still further, by coordinating co-management of chronic liver disease patients between hospitalists and hepatology. All three studies showed increased compliance with high-quality care such as early endoscopy for GI bleeding, early paracentesis, or prophylaxis for SBP; only Loftus showed a reduction in length of stay and mortality, but the other studies were not adequately powered. Redesigning workflows in this way has great potential to streamline necessary care, particularly in the time-sensitive inpatient setting.

Rigorous study design is an ongoing issue in quality improvement studies. The vast majority of studies included here utilized quasi-experimental design, most often a simple pretest/posttest approach. There were a handful of randomized and non-randomized controlled trials, including those on endoscopy retraining and improving patient bowel prep. It was unclear what factors led the authors to choose one design over another, although it is likely that financial and time restraints influenced the rigor of study design. Overall, the studies were heterogeneous and it was unclear

whether one design was more generally successful than others; however, the general argument can be made that the field of QI would benefit from the use of more rigorous designs using control groups such as RCTs. There were also a wide variation in sample size and no obvious pattern relating sample size to likelihood of having a successful intervention, which again may relate to overall study heterogeneity. Notably, none of the studies contained power calculations and it is likely that several of the smaller studies were underpowered. In general, there is a need for standardized application of research methods to quality improvement studies.

Our study is also notable for what was not found: we did not find any studies on improving patient-reported outcomes (PROs) or the impact of the patient experience on compliance with care. Since patient engagement and satisfaction are crucial factors in achieving high-quality care, explicit measurement of PROs should be an integral part of quality improvement. There was also a lack of studies on specific areas within GI such as IBS and chronic pancreatitis. We suspect this gap reflects the paucity of quality metrics for these diseases and the need for metric development.

Moving forward, a crucial question in quality improvement is how to focus limited resources on strategies likely to benefit the most patients. While technology can be a crucial and time-saving aid, it cannot replace dedicated staffing or reimagining of clinical workflows in certain cases. Here, we have attempted to characterize the attributes of successful QI programs for select disease areas in gastroenterology and consider the challenges involved in their implementation. Appropriate management of limited resources is becoming more important as we move away from volume-based to value-based health care, and accordingly QI will continue to play an important role in all fields of medicine. Value in this setting represents the health outcomes achieved per dollar spent and an alignment of patients' and providers' goals. Combining this approach with improving the patient experience should be the goal of quality improvement.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. OECD Health Statistics 2017. June 30th, 2017. Organization for Economic Co-operation and Development. <http://www.oecd.org/els/health-systems/health-data.htm>. Accessed August, 15, 2017.
2. Morgan L. US healthcare annual spending estimated to rise by 58% on average through 2024. *Am Health Drug Benefits*. 2015;8:272.
3. MacDorman MF, Mathews TJ, Mohangoo AD et al. International comparisons of infant mortality and related factors: United States and Europe. *Natl Cent Health Stat*. 2014;63(5):1–6.
4. Robeznieks A. US has the highest maternal death rate among developed countries. *Modern Healthcare*, May 6th, 2015. <http://www.modernhealthcare.com/article/20150506/NEWS/150509941>. Accessed August 15, 2017.
5. Xu J, Murphy SL, Kochanek KD et al. Mortality in the United States, 2015. *NCHS Data Brief*. 2016; No. 267.
6. Readmissions Reduction Program. Centers for Medicare and Medicaid Services. <https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html>. Accessed August 15, 2017.
7. Fonarow GC, Stevenson LW, Walden JA, et al. Impact of a comprehensive heart failure management program on hospital readmission and functional status of patients with advanced heart failure. *J Am Coll Cardiol*. 1997;30:725–732.
8. Krumholz HM, Amatruda J, Smith GL, et al. Randomized trial of an education and support intervention to prevent readmission of patients with heart failure. *J Am Coll Cardiol*. 2002;39:83–89.
9. Amarasingham R, Patel PC, Toto K, et al. Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective controlled study. *BMJ Qual Saf*. 2013;22912:998–1005.
10. Ong MK, Romano PS, Edington S, et al. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the better effectiveness after transition–heart failure (BEAT-HF) randomized clinical trial. *JAMA Intern Med*. 2016;22(12):310–318.
11. Hale TM, Jethwani K, Kandola MS, et al. A remote medication monitoring system for chronic heart failure patients to reduce readmissions: a two-arm randomized pilot study. *J Med Internet Res*. 2016;18:e91.
12. Atkinson S, Ingham J, Cheshire M, et al. Defining quality and quality improvement. *Clin Med (Lond)*. 2010;10:537–539.
13. Choosing Wisely: An Initiative of the ABIM Foundation. American Gastroenterology Association. <http://www.choosingwisely.org/societies/american-gastroenterological-association/>. Accessed January 5, 2018.
14. Physician Quality Reporting System 2015. American Gastroenterology Association. [https://www.gastro.org/practice/quality-initiatives/performance-measures/400-1708901-MHC-AGA\\_PQRS-IBD-Flashcard.pdf](https://www.gastro.org/practice/quality-initiatives/performance-measures/400-1708901-MHC-AGA_PQRS-IBD-Flashcard.pdf). Accessed January 5, 2018.
15. Huang Y, Don-Wauchope AC, Grey VL, et al. Improving serological test ordering patterns for the diagnosis of celiac disease through clinical laboratory audit of practice. *Clin Biochem*. 2012;45:455–459.
16. Tapper E, Finkelstein D, Mittleman M, et al. A quality improvement initiative decreases 30-day readmission rates for patients with cirrhosis: a prospective trial. *Clin Gastroenterol Hepatol*. 2016;14:753–759.
17. Ghaoui R, Friderici J, Desilets DJ, et al. Outcomes associated with a mandatory gastroenterology consultation to improve the quality of care of patients hospitalized with decompensated cirrhosis. *J Hosp Med*. 2015;10:236–241.
18. Desai AP, Satoskar R, Appannagari A, et al. Co-management between hospitalist and hepatologist improves the quality of care of inpatients with chronic liver disease. *J Clin Gastroenterol*. 2014;48:e30–e36.
19. Kennedy NA, Rodgers A, Altus R, et al. Optimisation of hepatocellular carcinoma surveillance in patients with viral hepatitis: a quality improvement study. *Intern Med J*. 2013;43:772–777.
20. Aberra FB, Essenmacher M, Fisher N, et al. Quality improvement measures lead to higher surveillance rates for hepatocellular carcinoma in patients with cirrhosis. *Dig Dis Sci*. 2013;58:1157–1160. <https://doi.org/10.1007/s10620-012-2461-4>.
21. Loy V, Kwiatk J, Dodda A. Performance feedback improves compliance with quality measures. *Am J Med Qual*. 2016;31:118–124.
22. Inra J, Naylor J, Rosenblatt M, et al. Comparison of colonoscopy quality measures across various practice settings and the impact

- of performance scorecards. *Dig Dis Sci*. 2017;62:894–902. <https://doi.org/10.1007/s10620-016-4410-0>.
23. Shaukat A, Oancea C, Bond J, et al. Variation in detection of adenomas and polyps by colonoscopy and change over time with a performance improvement program. *Clin Gastroenterol Hepatol*. 2009;7:1335–1340.
  24. Clayton D, Palmer W, Robison S, et al. Colonoscopy bowel preparation quality improvement for patients with decompensated cirrhosis undergoing evaluation for liver transplantation. *Clin Transplant*. 2016;30:1236–1241.
  25. Kahi C, Ballard D, Shah A, et al. Impact of a quarterly report card on colonoscopy quality measures. *Gastrointest Endosc*. 2013;77:925–931.
  26. Mai HD, Sanowski RA, Waring JP. Improved patient care using the A/S/G/E guidelines on quality assurance: a prospective comparative study. *Gastrointest Endosc*. 1991;37:597–599.
  27. Abdul-Baki H, Schoen R, Dean K, et al. Public reporting of colonoscopy quality is associated with an increase in endoscopist adenoma detection rate. *Gastrointest Endosc*. 2015;82:676–682.
  28. Ross W, Thirumurthi S, Lynch P, et al. Detection rates of pre-malignant polyps during screening colonoscopy: time to revise quality standards? *Gastrointest Endosc*. 2015;81:567–574.
  29. Coe S, Crook J, Diehl N, et al. An endoscopic quality improvement program improves detection of colorectal adenomas. *Am J Gastroenterol*. 2013;108:219–226.
  30. Ussui V, Coe S, Rizk C, et al. Stability of increased adenoma detection at colonoscopy. Follow-up of an endoscopic quality improvement program-EQUIP-II. *Am J Gastroenterol*. 2015;110:489–496.
  31. Wallace MB, Crook JE, Thomas CS, et al. Effect of an endoscopic quality improvement program on adenoma detection rates: a multicenter cluster-randomized controlled trial in a clinical practice setting (EQUIP-3). *Gastrointest Endosc*. 2017;85:538–545.
  32. Rajasekhar P, Rees C, Bramble M, et al. A multicenter pragmatic study of an evidence based intervention to improve adenoma detection: the Quality Improvement in Colonoscopy (QIC) study. *Endoscopy*. 2015;47(3):217–224.
  33. Imperiali G, Minoli G, Meucci M, et al. Effectiveness of a continuous quality improvement program on colonoscopy practice. *Endoscopy*. 2007;39:314–318.
  34. Barclay R, Vicari J, Greenlaw R, et al. Effect of a time-dependent colonoscopic withdrawal protocol on adenoma detection during screening colonoscopy. *Clin Gastroenterol Hepatol*. 2008;6:1091–1098.
  35. Hayat U, Lee P, Lopez R, et al. Online educational video improves bowel preparation and reduces the need for repeat colonoscopy within three years. *Am J Med*. 2016;129:1219.
  36. Park J, Kim M, Kim H, et al. A randomized controlled trial of an educational video to improve quality of bowel preparation for colonoscopy. *BMC Gastroenterol*. 2016;16:64.
  37. Hseuh F, Wang H, Sun C, et al. The effect of different patient education methods on quality of bowel cleanliness in outpatients receiving colonoscopy examination. *App Nurs Res*. 2014;27:e1–e5.
  38. Grassini M, Verna C, Battaglia E, et al. Education improves colonoscopy appropriateness. *Gastrointest Endosc*. 2008;67:88–93.
  39. Player MS, Gill JM, Mainous AG, et al. An electronic medical record-based intervention to improve quality of care for gastroesophageal reflux disease (GERD) and atypical presentations of GERD. *Qual Prim Care*. 2010;18:223–229.
  40. Walsh K, Kwan D, Marr P, et al. Deprescribing in a family health team: a study of chronic proton pump inhibitor use. *J Prim Health Care*. 2016;8:164–171.
  41. Loftus TJ, Go KL, Hughes SJ, et al. Improved outcomes following implementation of an acute gastrointestinal bleeding multidisciplinary protocol. *J Trauma Acute Care Surg*. 2017;83:41–46.
  42. Johnson EA, Spier BJ, Leff JA, et al. Optimising the care of patients with cirrhosis and gastrointestinal haemorrhage: a quality improvement study. *Aliment Pharmacol Ther*. 2011;34:76–82.
  43. Pfau P, Cooper G, Carlson M, et al. Success and shortcomings of a clinical care pathway in the management of acute non-variceal upper gastrointestinal bleeding. *Am J Gastroenterol*. 2004;99:425–431.
  44. 2015 National Healthcare Quality and Disparities Report and 5th Anniversary Update on the National Quality Strategy. Agency for Healthcare Research and Quality. <https://www.ahrq.gov/research/findings/nhqrdr/nhqrdr15/priorities.html>. Accessed January 5, 2018.
  45. Sapir T, Moreo K, Carter JD, et al. Continuing medical education improves gastroenterologists' compliance with inflammatory bowel disease quality measures. *Dig Dis Sci*. 2016;61:1862–1869. <https://doi.org/10.1007/s10620-016-4061-1>.
  46. Greene L, Moreo K. Quality improvement education to improve performance on ulcerative colitis quality measures and care processes aligned with National Quality Strategy priorities. *BMJ Qual Improv Rep*. 2015;4:u208829-w3554.
  47. Greene L, Sapir T, Moreo K, et al. Impact of quality improvement educational interventions on documented adherence to quality measures for adults with Crohn's disease. *Inflamm Bowel Dis*. 2015;21:2165–2171.
  48. Walsh AJ, Wetman M, Burger D, et al. Implementing guidelines on the prevention of opportunistic infections in inflammatory bowel disease. *J Crohns Colitis*. 2013;7:e449–e456.
  49. Baysari MT, Tarig A, Day RO. Alert override as a habitual behavior—a new perspective on a persistent problem. *J Am Med Inform Assoc*. 2017;24:409–412.