Published in final edited form as:

Int J Med Inform. 2015 December ; 84(12): 1111–1117. doi:10.1016/j.ijmedinf.2015.07.006.

Computer versus physician identification of gastrointestinal alarm features

Christopher V. Almario[a,b,c,d], **William D. Chey**[e], **Sentia Iriana**[a], **Francis Dailey**[a], **Karen Robbins**[a], **Anish V. Patel**[a], **Mark Reid**[b,d], **Cynthia Whitman**[a,d], **Garth Fuller**[a,d], **Roger Bolus**[d], **Buddy Dennis**[f], **Rey Encarnacion**[f], **Bibiana Martinez**[a,b,d], **Jennifer Soares**[a,b,d], **Rushaba Modi**[a,b,c], **Nikhil Agarwal**[a,b,c], **Aaron Lee**[a], **Scott Kubomoto**[a], **Gobind Sharma**[a], **Sally Bolus**[d], **Lin Chang**[c], and **Brennan M.R. Spiegel**[a,b,d,*]

[a] Division of Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, CA, USA

[b] Division of Gastroenterology, VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA

[c] Division of Digestive Diseases, UCLA, Los Angeles, CA, USA

[d] Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Los Angeles, CA, USA

[e] Division of Gastroenterology, University of Michigan, Ann Arbor, MI, USA

[f] UCLA Computing Technology Research Laboratory (CTRL), Los Angeles, CA, USA

[*] Corresponding author at: Brennan M.R. Spiegel, MD, MSHS, Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Director of Health Services Research, Cedars-Sinai Health System, Professor-in-Residence of Medicine and Public Health, 116 No. Robertson Blvd., 4th Floor, Los Angeles, California 90048 United States. Brennan.Spiegel@cshs.org.

Authors contribution

Christopher V. Almario, MD, MSHPM: Planning and conducting the study, collecting and interpreting data, drafting the manuscript, approval of final draft submitted.
William D. Chey, MD: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted.
Sentia Iriana, MD: Planning and conducting the study, approval of final draft submitted.
Francis Dailey, MD: Planning and conducting the study, approval of final draft submitted.
Karen Robbins, MD: Planning and conducting the study, approval of final draft submitted.
Anish V. Patel, MD: Planning and conducting the study, approval of final draft submitted.
Mark Reid, PhD: Interpreting data, drafting the manuscript, approval of final draft submitted.
Cynthia Whitman, MPH: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted.
Garth Fuller, MS: Interpreting data, drafting the manuscript, approval of final draft submitted.
Roger Bolus, PhD: Planning and conducting the study, interpreting data, approval of final draft submitted.
Buddy Dennis, PhD: Planning and conducting the study, approval of final draft submitted.
Rey Encarnacion, BS: Planning and conducting the study, approval of final draft submitted.
Bibiana Martinez, MPH: Planning and conducting the study, approval of final draft submitted.
Jennifer Soares, MSPH: Planning and conducting the study, approval of final draft submitted.
Rushaba Modi, MD: Planning and conducting the study, approval of final draft submitted.
Nikhil Agarwal, MD: Planning and conducting the study, approval of final draft submitted.
Aaron Lee, MD: Planning and conducting the study, approval of final draft submitted.
Scott Kubomoto, MD: Planning and conducting the study, approval of final draft submitted.
Gobind Sharma, MD: Planning and conducting the study, approval of final draft submitted.
Sally Bolus, MS: Planning and conducting the study, approval of final draft submitted.
Lin Chang, MD: Planning and conducting the study, interpreting data, approval of final draft submitted.
Brennan M.R. Spiegel, MD, MSHS: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted.

Conflict of interest

Drs. Chey and Spiegel are principals in My Total Health. The remaining authors do not have any disclosures.

Abstract

Objective—It is important for clinicians to inquire about “alarm features” as it may identify those at risk for organic disease and who require additional diagnostic workup. We developed a computer algorithm called Automated Evaluation of Gastrointestinal Symptoms (AEGIS) that systematically collects patient gastrointestinal (GI) symptoms and alarm features, and then “translates” the information into a history of present illness (HPI). Our study's objective was to compare the number of alarms documented by physicians during usual care vs. that collected by AEGIS.

Methods—We performed a cross-sectional study with a paired sample design among patients visiting adult GI clinics. Participants first received usual care by their physicians and then completed AEGIS. Each individual thus contributed both a physician-documented and computer-generated HPI. Blinded physician reviewers enumerated the positive alarm features (hematochezia, melena, hematemesis, unintentional weight loss, decreased appetite, and fevers) mentioned in each HPI. We compared the number of documented alarms within patient using the Wilcoxon signed-rank test.

Results—Seventy-five patients had both physician and AEGIS HPIs. AEGIS identified more patients with positive alarm features compared to physicians (53% vs. 27%; $p < .001$). AEGIS also documented more positive alarms (median 1, interquartile range [IQR] 0–2) vs. physicians (median 0, IQR 0–1; $p < .001$). Moreover, clinicians documented only 30% of the positive alarms self-reported by patients through AEGIS.

Conclusions—Physicians documented less than one-third of red flags reported by patients through a computer algorithm. These data indicate that physicians may under report alarm features and that computerized “checklists” could complement standard HPIs to bolster clinical care.

Keywords

Alarm features; Checklists; Patient-provider portal

1. Introduction

Adoption of electronic health records (EHRs) has proceeded at an accelerated rate in large part due to the Health Information Technology for Economic and Clinical Health Act [1]. This has fundamentally changed the way healthcare providers document, monitor, and share information [2]. With respect to documentation, EHRs allow for improved availability of prior clinical notes, more streamlined data organization, and eliminates issues related to illegibility [3,4]. This helps narrow the gap between information and action that at times can result in delayed or inadequate care [5]. Prior reports have also found that EHRs improved the quality of physician notes documented in the medical record. Using a validated instrument for measuring the quality of EHR clinical notes, Burke and colleagues found that introduction of an EHR led to higher quality scores across all core and non-core elements of the clinical notes for patients with type II diabetes [6,7]. Similarly, Roshanov et al. noted that an electronic, diabetes-specific, chronic disease management system captured more clinically important data versus dictated notes [8].

Our group, in an effort to improve the quality and comprehensiveness of gastrointestinal (GI) notes in EHR-integrated practices, developed a computer algorithm called Automated Evaluation of Gastrointestinal Symptoms (AEGIS) that is available through a patient-provider portal [9]. AEGIS systematically collects both GI symptom and alarm feature information from patients through its history of present illness (HPI) and alarm feature modules, respectively. We previously compared the quality of reports generated by the AEGIS HPI module to those written by physicians during usual care in academic GI clinics [9]. Here, blinded raters deemed the computer-generated HPIs to be more complete, succinct and useful compared to physician-documented HPIs.

While EHRs have been shown on balance to improve note quality, little is known about the effectiveness of GI providers at documenting alarm features within an EHR-integrated practice. Identification of alarm features or “red flags” remains an important part of the medical interview. Current national guidelines for common GI disorders including gastroesophageal reflux disease, dyspepsia, irritable bowel syndrome (IBS), among others, all recommend assessing for alarm features [10–15]. More importantly, it may identify those with organic disease and who require additional diagnostic evaluation. Patel and colleagues found that IBS patients with alarm features were significantly more likely to have organic GI disease versus those without such features [16]. Because of the importance of assessing for alarm features, we performed a cross-sectional study with a paired sample design to compare the number of alarm features documented in physician HPIs composed during usual care versus that captured by AEGIS’ alarm features module.

2. Materials and methods

2.1. Study overview

We compared the number of alarm features documented in HPIs generated through two methods on the same patients: (1) physician HPIs documented in the EHR; and (2) computer-generated HPIs created by a computer algorithm designed to systematically collect patient information. Blinded physicians without knowledge about the purpose of the study enumerated the number of alarm features in the two sets of HPIs. We conducted the study in GI clinics at the University of California, Los Angeles (UCLA) and the West Los Angeles Veterans Affairs (WLAVA) Medical Center. The UCLA (IRB #13-000337) and WLAVA (IRB PCC #2014-020138) Institutional Review Boards approved this study.

2.2. Automated Evaluation of Gastrointestinal Symptoms (AEGIS)

Our groups at Cedars-Sinai Medical Center and the University of Michigan developed a computer algorithm called AEGIS that is available through a patient-provider portal called My GI Health (www.MyGIHealth.org). We describe the AEGIS algorithm in detail elsewhere [9]. Briefly, we created AEGIS with the goal of improving the patient-physician relationship. Healthcare delivery currently is largely centered around the clinic visit, yet patients spend the vast majority of their time outside of the exam room. Based on this, we designed the AEGIS algorithm to systematically collect information from the patient prior to the visit, allowing them to tell their own “story” about their GI symptoms without time or locale restraints. AEGIS takes this information and creates a symptom report that includes a

GI symptom “heat map” [9] and GI HPI (detailed below) that physicians can review prior to or concurrent with seeing the patient; this leaves more time during the visit for counseling and education which we hypothesize positively impacts patient–physician communication. Moreover, AEGIS supports both the patient and physician by creating a tailored educational prescription informed by the information previously collected by the algorithm. Formal studies testing whether AEGIS improves patient outcomes versus usual care are currently underway.

The AEGIS algorithm contains an HPI module and an alarm features module. Patients are first directed to the HPI module, which employs computerized adaptive testing to guide them through National Institutes of Health (NIH) GI Patient Reported Outcomes Measurement Information System (PROMIS®) questionnaires [17] as well as questions drawn from a library of over 300 symptom attributes measuring the timing, severity, frequency, location, quality, bother, and character of their GI symptoms (abdominal pain, bloat/gas, diarrhea, constipation, bowel incontinence, heartburn/reflux, dysphagia, and nausea/vomiting). Once the questions are completed, the information is transformed into a full narrative HPI (Fig. 1, Paragraphs 1–2) written in language familiar to clinicians.

After completing the HPI module, AEGIS then guides patients through the alarm features module. Here, AEGIS systematically inquires about presence of alarm features including hematochezia, melena, hematemesis, unintentional weight loss, decreased appetite, and fevers (Appendix Fig. A1 displays the questions included in AEGIS). These alarm symptoms were chosen based on clinical guidelines [10–12,14,15] and input from experienced gastroenterologists (W.D.C., L.C., and B.M.R.S.). This information is subsequently added to the HPI (Fig. 1, Paragraph 3). The alarm features paragraph details presence and absence of alarm features, including information about their onset, frequency, and amount, when applicable. Table 1 provides a list of all alarm features and attributes in the AEGIS HPI.

2.3. Setting

We performed a cross-sectional study with a paired sample design among patients who visited GI clinics at WLAVA or UCLA. The WLAVA GI clinic is an academic teaching practice staffed by GI attending physicians, GI fellows, internal medicine residents, and GI physician assistants. Notes generated during the encounter were typed directly into the Veterans Affairs Computerized Patient Record System. Attending physicians primarily staffed the GI clinics at Ronald Reagan UCLA Medical Center and Santa Monica-UCLA Medical Center. Here, providers either typed or dictated notes into the EHR developed by Epic Systems Corporation (Verona, WI).

2.4. Patients

We enrolled patients aged 18 years or older who had one or more active GI symptoms at the time of their visit. Patients were required to read and write English and have basic computing skills. We excluded asymptomatic patients, or those seeking consultation only for abnormal blood tests, cancer screening, or other non-symptomatic indications.

All patients first received usual care in their respective GI clinics before completing AEGIS. While AEGIS reports would be more useful before a visit in clinical practice, administering AEGIS prior to the visit would potentially bias the physician encounter by priming the patient and unblinding the study. Patients from the WLAVA GI clinic were seen between July and December 2013, and were recruited either directly in clinic or through mailed recruitment materials. During clinic, patients who met inclusion criteria were invited to complete AEGIS on a computer after completion of their physician visit. For the remaining patients not recruited in clinic, we reviewed their EHR charts the following day, and those meeting inclusion criteria were sent a letter inviting them to complete AEGIS online. UCLA patients were recruited between January and March 2014 solely through mailed recruitments.

2.5. Primary outcome

Our primary outcome was number of positive alarm features documented in the AEGIS versus physician HPIs. For the purposes of this study, we included the review of systems section from the physician notes to the physician HPI because some providers record alarm features within that section.

Two blinded physician reviewers not associated with the study team independently identified alarm features mentioned in each report. The reviewers were not informed of the purpose of the study or that a computer created half of the HPIs. The order of the AEGIS and physician-documented HPIs was randomized prior to sending them to the reviewers. We evaluated inter-rater reliability between reviewers using a kappa statistic for each alarm feature. In cases of rater disagreement, a third, blinded physician rater made a final determination.

2.6. Secondary outcomes

We assessed the proportion of patient self-reported (through AEGIS) positive alarm features that were captured by physicians, and vice versa. Moreover, for each documented alarm, we performed an analysis comparing the number of attributes (onset, frequency, and amount) documented in the physician vs. AEGIS HPIs. Another blinded physician reviewer not associated with the study team evaluated each HPI that documented at least one positive alarm feature. For each positive alarm, the reviewer determined whether the HPI provided information regarding the alarm feature's onset, frequency (not applicable for melena, weight loss, decreased appetite, or fevers), and amount (not applicable for melena, decreased appetite, or fevers).

2.7. Statistical analysis

Statistical analyses were performed using Stata 13.1 (StataCorp LP, College Station, TX). A two-tailed *p*-value of less than .05 was considered significant in all analyses. For bivariate analyses, we used either the Wilcoxon rank-sum test or Kruskal–Wallis test to compare groups. To compare proportions between groups, we used the two-sample test of proportions.

The primary outcome was difference in number of alarm features documented in computer-generated versus physician HPIs. We first performed bivariate analyses looking at the association between number of alarm features and potential confounding factors such as patient age (at time of GI clinic visit), sex, race/ethnicity, physician HPI author (attending, fellow, resident/physician assistant), HPI input method (typed or dictated), visit type (initial or follow-up), and site of care. No significant associations were found (all $p > .05$), so we opted to compare the number of documented alarm features in the computer-generated versus physician HPIs using the Wilcoxon signed-rank test. We also generated a scatterplot and calculated R^2 of the ordinary least squares regression line.

3. Results

3.1. Patients and physicians

Table 2 presents information regarding the 75 study patients. Each patient contributed two HPIs leading to a total of 150 HPIs. Twenty-nine patients were directly recruited in the WLAVA GI clinic while the remaining 46 patients were enrolled through mailed recruitment materials following their visit. Among those recruited by mail, 8% (18/220) of WLAVA patients and 7% (28/428) of UCLA patients completed AEGIS a median of seven days (range: 2–29 days) after their visit. The 75 physician HPIs were written by 23 unique providers (six GI attending physicians, nine GI sub-specialty fellows, seven internal medicine residents, and one GI physician assistant). GI attendings authored 35% (26/75) of the HPIs while GI fellows (48%, 36/75), internal medicine residents (13%, 10/75), and GI physician assistants (4%, 3/75) were responsible for the remaining ones. Ninety-two percent (69/75) of the physician HPIs were typed directly into the EHR and 8% (6/75) were verbally dictated.

3.2. Alarm features inter-rater agreement

The alarm feature assessments by the two independent blinded physician reviewers matched for 93% (140/150) of the HPIs. Appendix Table A1 shows the kappa statistics according to HPI source and individual alarm features. For the physician HPIs, the inter-rater agreement ranged from moderate to near perfect across the various alarm features. For AEGIS HPIs, there was near perfect agreement across alarms.

3.3. Primary analysis – comparing number of alarm features in AEGIS vs. physician HPIs

Overall, AEGIS identified more patients reporting one or more positive alarm features vs. physicians (53% (40/75) vs. 27% (20/75); $p < .001$). Table 3 presents the number of documented alarm features according to HPI source. AEGIS HPIs documented significantly more positive alarm features compared to physician-generated HPIs.

A scatterplot comparing the number of alarm features documented in the two sets of HPIs can be found in Fig. 2. The ordinary least square result had a lower than expected correlation, indicating that AEGIS captured more positive alarm features, on average, than physician HPIs.

3.4. Subgroup analyses

When focusing only on initial visits (n = 39), AEGIS HPIs still had significantly more documented positive alarm features versus physician HPIs (Table 3). Because alarm symptoms could potentially change over short periods of time, we also performed a subgroup analysis that only included individuals who completed AEGIS within one week of their clinic visit (*n* = 55). Here, AEGIS HPIs again documented more positive alarm features (Table 3).

3.5. Secondary analyses

3.5.1. Alarm feature matching assessment—We evaluated whether the positive alarm features documented in the physician HPI matched the patients' self-reported alarms in AEGIS. Overall, physicians documented 30% (21/71) of the positive alarm features reported by patients through AEGIS. Table 4 lists this analysis according to individual alarm features. Results were largely unchanged when limiting this to patients who either were presenting for an initial visit (38%, 14/37) or who completed AEGIS within 1 week of their clinic visits (31%, 16/51).

We also performed an analysis determining whether physicians documented positive alarm features that were not captured by AEGIS. Overall, AEGIS identified 78% (21/27) of the alarms documented by physicians. Findings were similar when limiting the analysis to patients presenting for an initial visit (74%, 14/19) and those completing AEGIS within 1 week of their visit (76%, 16/21).

3.5.2. Alarm feature attribute assessment—Appendix Table 2 depicts the number of positive alarm attributes documented in the physician and AEGIS HPIs. Physician HPIs provided less detail for hematochezia (median 1, interquar-tile range [IQR] 1–2 vs. median 3, IQR 3–3; $p < .001$) and decreased appetite (median 1, IQR 0–1 vs. median 1, IQR 1–1; $p = .05$) when compared to AEGIS HPIs. No difference was seen for hematemesis and unintentional weight loss when comparing the groups. We were unable to perform this analysis for melena and fevers because of limited number of cases.

4. Discussion

We found that a computer algorithm delivered through a patient-provider portal identified more alarm features vs. physicians during usual care in academic GI clinics. These data indicate that GI physicians may underreport alarm features. The AEGIS algorithm tested in this study offers one model for how a computerized alarm feature "checklist" could serve as a vital complement to standard HPIs in EHR-integrated practices.

There are many potential explanations for why physicians documented fewer alarm features in this study. Healthcare providers now practice in a digital era that has led to information and data overload; this may have led some providers to overlook common tasks such as asking about alarm features. Providers are also responsible for an ever-increasing number of clinical and administrative tasks, and they may not have had the time to systematically inquire about alarm features in every patient. Similarly, in this pressured environment, physicians may have opted to focus their documentation efforts on other core sections of the

note. Employing AEGIS pre-visit (as it is intended to be used in clinical practice), provides a potential solution to these issues as information regarding alarm features is systematically collected from the patient prior to the clinic visit. AEGIS then creates an alarm features report with potentially actionable information that effectively functions as a “checklist” supporting the physician. “Checklists” have long been used in both medical and surgical fields, and have been shown to improve communication and reduce complications, morbidity and mortality [18,19]. They have also been used successfully in other high-risk fields like aviation and engineering [20]. A limitation of the current study is that we could not use the checklists prior to the visit because it would have biased the physician's “status quo” and typical history-taking by pre-identifying alarm features; i.e., we could not have directly compared side-by-side performance.

AEGIS also goes beyond just detecting presence or absence of alarm features, as it collects additional detail including onset, frequency, and amount. We found that AEGIS HPIs, compared to physician HPIs, provided more detailed assessments for hematochezia and lack of appetite. This provides physicians with clinically meaningful information and lifts some of the burden of data collection and documentation off of providers.

Similar to our previous report comparing physician-documented versus patient self-reported chief complaint, [9] we discovered a frequent mismatch between physician HPI-documented and patient self-reported alarm features. From the patient perspective, it is possible that some patients did not report alarm features to their physician or felt more comfortable reporting it to the computer. Lucas and colleagues found that patients are comfortable disclosing health information to “virtual human” interviewers as supportive and “safe” interaction partners [21]. The mismatch may also have been compounded by the fact that patients completed AEGIS after their GI consultation; the visit may have “primed” patients to select more alarm features in AEGIS than were discussed during the visit. Conversely, from the provider side, some physicians may have inquired about certain alarm features during the visit, but did not document the presence or absence of such symptoms. Physicians also may not have explicitly asked their patients about alarm features that were not pertinent to the chief complaint. Yet, even if an alarm feature is not related to the chief complaint, it may still be clinically relevant and important to address. While our study design limits our ability to determine why the mismatch exists, it is worth reinforcing that patients in real clinical practice are meant to complete AEGIS before their GI consultation. Completion of AEGIS pre-visit and having the report available to the physician during the visit might improve upon this mismatch and enhance patient-centered care, but that must be formally investigated.

However, we should note that just because AEGIS collects more alarm feature information compared to physicians does not mean that computers could ever replace healthcare providers. For instance, while AEGIS identified the majority of physician-documented alarms, some were missed. This reinforces the point that the AEGIS alarm feature “checklist” is a tool that supports physicians’ decision making processes by collecting and organizing information. The art of medicine still requires clinicians to interpret such data as well as to gather any additional clinically relevant information during the course of the patient encounter. We should also acknowledge the risk that AEGIS may “over-diagnose”

alarm features potentially leading to unnecessary invasive testing. However, the AEGIS "checklist" again is a supportive tool and it is incumbent on the physician to determine whether an alarm feature is clinically relevant by factoring in the patient's history, physical exam, and prior diagnostic workup.

We should also note that it is not known whether improved identification of red flags leads to better outcomes. Yet, current national guidelines recommend routine assessment of alarm features for many common GI disorders and it remains an important part of the medical interview [10–15]. While alarm features often have poor sensitivity and specificity for diagnosing gastrointestinal cancers [22–24], Patel and colleagues discovered that alarm features have utility in predicting non-malignant GI disorders [16]. Namely, they found that 28% of IBS patients with alarm features had organic GI disease compared to only 15% among those without such features ($p = .002$). The non-malignant diagnoses seen among IBS patients with alarm features were ulcerative colitis, Crohn's disease, microscopic colitis and Celiac disease, all of which are treatable conditions. While these results cannot be extrapolated to all GI disease states, it strongly suggests that assessing for alarm features has clinical utility. Moreover, in the absence of more precise clinical markers of severe organic disease, physicians must rely on soliciting and documenting alarm features to produce a full assessment and plan.

This study has limitations. First, our study design did not allow us to determine accuracy of the presence of alarm features, as we did not record the patient encounters. Yet, the AEGIS HPI was entirely based on patient self-report, and patients are always the first source – by definition – of patient-reported signs and symptoms. Second, it is possible that the reviewers might have detected a pattern indicating structural similarities between the computer HPIs. The reviewers, however, were not aware of the purpose of the study or of AEGIS and its capabilities. Moreover, our primary outcome of documentation of alarm features in the HPI was an objective and verifiable finding. Third, patients completed AEGIS after rather than before their visit. While AEGIS should optimally be administered before the visit, this was not feasible as this would prime the patient and potentially unblind the study, as previously emphasized.

In summary, this study demonstrates that physicians may under report alarm features in academic GI clinics. AEGIS provides a model for how a computer algorithm can generate an alarm feature "checklist" that complements standard HPIs in an EHR-integrated practice. Future research determining whether better identification of alarm features leads to improved clinical outcomes are needed.

Acknowledgements

NIH/NIAMS U01 AR057936A, the National Institutes of Health through the NIH Roadmap for Medical Research Grant (AR052177). Dr. Almario was supported by a National Institutes of Health T32 training grant (NIH T32DK07180-40) during his gastroenterology and health services research training at UCLA. Support for the My GI Health portal used in this study was obtained from Ironwood Pharmaceuticals. The above funding sources did not play any role in the conduct of this study.

Appendix A

See Tables A1 and A2 and Fig. A1.

Table A1

Alarm feature inter-rater agreement.

Source of HPI and positive alarm features	Kappa
Physician HPIs:	
Hematochezia	0.79
Melena	0.66
Hematemesis	0.49
Unintentional weight loss	0.85
Decreased appetite	1.0
Fevers	_ [a]
AEGIS HPIs:	
Hematochezia	1.0
Melena	1.0
Hematemesis	1.0
Unintentional weight loss	1.0
Decreased appetite	0.94
Fevers	1.0

AEGIS, Automated Evaluation of Gastrointestinal Symptoms; HPI, history of present illness.

[a] Not enough cases to calculate kappa statistic.

Table A2

Alarm feature attribute assessment.

Alarm feature	Num. of alarm feature attributes		*p*-value [*]
	Physician HPI	AEGIS HPI	
Hematochezia [a]	1 [1–2] (*n* = 6)	3 [3–3] (*n* = 9)	<.001
Hematemesis [a]	2.5 [2–3] (*n* = 2)	3 [3–3] (*n* = 4)	.16
Unintentional weight loss [b]	2 [2–2] (*n* = 11)	2 [2–2] (*n* = 22)	.19
Decreased appetite [c]	1 [0–1] (*n* = 7)	1 [1–1] (*n* = 24)	.05

Data are presented as median [interquartile range].

Unable to perform analysis for melena and fevers because of limited number of cases.

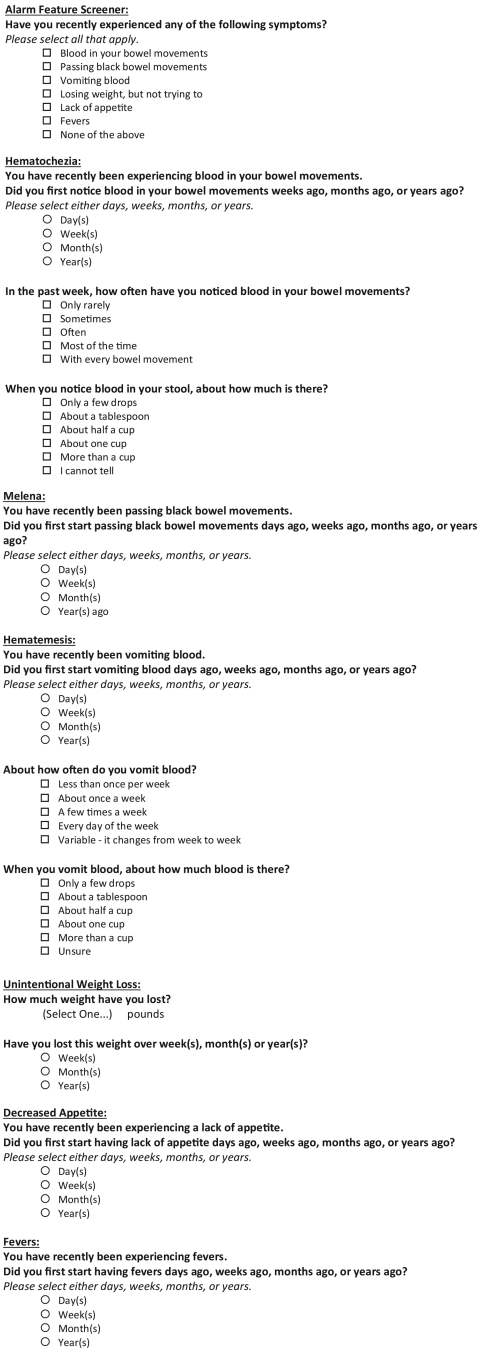
AEGIS, Automated Evaluation of Gastrointestinal Symptoms; HPI, history of present illness.

[*] Wilcoxon rank-sum test.

[a] Assessed for description of onset, frequency, and amount.

[b] Assessed for description of onset and amount.

[c] Assessed for description of onset.



Alarm Feature Screener:
Have you recently experienced any of the following symptoms?
Please select all that apply.
- ☐ Blood in your bowel movements
- ☐ Passing black bowel movements
- ☐ Vomiting blood
- ☐ Losing weight, but not trying to
- ☐ Lack of appetite
- ☐ Fevers
- ☐ None of the above

Hematochezia:
You have recently been experiencing blood in your bowel movements.
Did you first notice blood in your bowel movements weeks ago, months ago, or years ago?
Please select either days, weeks, months, or years.
- ○ Day(s)
- ○ Week(s)
- ○ Month(s)
- ○ Year(s)

In the past week, how often have you noticed blood in your bowel movements?
- ☐ Only rarely
- ☐ Sometimes
- ☐ Often
- ☐ Most of the time
- ☐ With every bowel movement

When you notice blood in your stool, about how much is there?
- ☐ Only a few drops
- ☐ About a tablespoon
- ☐ About half a cup
- ☐ About one cup
- ☐ More than a cup
- ☐ I cannot tell

Melena:
You have recently been passing black bowel movements.
Did you first start passing black bowel movements days ago, weeks ago, months ago, or years ago?
Please select either days, weeks, months, or years.
- ○ Day(s)
- ○ Week(s)
- ○ Month(s)
- ○ Year(s) ago

Hematemesis:
You have recently been vomiting blood.
Did you first start vomiting blood days ago, weeks ago, months ago, or years ago?
Please select either days, weeks, months, or years.
- ○ Day(s)
- ○ Week(s)
- ○ Month(s)
- ○ Year(s)

About how often do you vomit blood?
- ☐ Less than once per week
- ☐ About once a week
- ☐ A few times a week
- ☐ Every day of the week
- ☐ Variable - it changes from week to week

When you vomit blood, about how much blood is there?
- ☐ Only a few drops
- ☐ About a tablespoon
- ☐ About half a cup
- ☐ About one cup
- ☐ More than a cup
- ☐ Unsure

Unintentional Weight Loss:
How much weight have you lost?
(Select One...) pounds

Have you lost this weight over week(s), month(s) or year(s)?
- ○ Week(s)
- ○ Month(s)
- ○ Year(s)

Decreased Appetite:
You have recently been experiencing a lack of appetite.
Did you first start having lack of appetite days ago, weeks ago, months ago, or years ago?
Please select either days, weeks, months, or years.
- ○ Day(s)
- ○ Week(s)
- ○ Month(s)
- ○ Year(s)

Fevers:
You have recently been experiencing fevers.
Did you first start having fevers days ago, weeks ago, months ago, or years ago?
Please select either days, weeks, months, or years.
- ○ Day(s)
- ○ Week(s)
- ○ Month(s)
- ○ Year(s)

Fig. A1.
Alarm feature questions included in Automated Evaluation of Gastrointestinal Symptoms. Note: ☐ = checkbox; O = leads to free text box where patients can enter the number of days, weeks, months or years since the onset of the alarm feature.

Abbreviations

AEGIS	Automated Evaluation of Gastrointestinal Symptoms

EHR	electronic health record
GI	gastrointestinal
HPI	history of present illness
IBS	irritable bowel syndrome
IQR	interquartile range
NIH	National Institutes of Health
PROMIS®	Patient Reported Outcome Measurement Information System
UCLA	University of California, Los Angeles
WLAVA	West Los Angeles Veterans Affairs

References

1. Blumenthal D. Launching HITECH. N. Engl. J. Med. 2010; 362:382–385. [PubMed: 20042745]
2. DesRoches CM, Charles D, Furukawa MF, et al. Adoption of electronic health records grows rapidly, but fewer than half of US hospitals had at least a basic system in 2012. Health Aff. (Millwood). 2013; 32:1478–1485. [PubMed: 23840052]
3. Miller RH, Sim I. Physicians' use of electronic medical records: barriers and solutions. Health Aff. (Millwood). 2004; 23:116–126. [PubMed: 15046136]
4. Embi PJ, Yackel TR, Logan JR, et al. Impacts of computerized physician documentation in a teaching hospital: perceptions of faculty and resident physicians. J. Am. Med. Inform. Assoc. 2004; 11:300–309. [PubMed: 15064287]
5. Handler, T.; Holtmeier, R.; Metzger, J., et al. HIMSS electronic health record definitional model version 1.0. The Healthcare Information and Management Systems Society (HIMSS) Electronic Health Record Committee; 2003.
6. Burke HB, Hoang A, Becher D, et al. QNOTE: an instrument for measuring the quality of EHR clinical notes. J. Am. Med. Inform. Assoc. 2014; 21:910–916. [PubMed: 24384231]
7. Burke HB, Sessums LL, Hoang A, et al. Electronic health records improve clinical note quality. J. Am. Med. Inform. Assoc. 2015; 22:199–205. [PubMed: 25342178]
8. Roshanov PS, Gerstein HC, Hunt DL, et al. Impact of a computerized system for evidence-based diabetes care on completeness of records: a before-after study. BMC Med. Inform. Decis. Mak. 2012; 12:63. [PubMed: 22769425]
9. Almario CV, Chey W, Kaung A, et al. Computer-generated vs. physician-documented history of present illness (HPI): results of a blinded comparison. Am. J. Gastroenterol. 2015; 110:170–179. [PubMed: 25461620]
10. American Gastroenterological Association medical position statement: irritable bowel syndrome. Gastroenterology. 2002; 123:2105–7. [PubMed: 12454865]
11. Bharucha AE, Dorn SD, Lembo A, et al. American Gastroenterological Association medical position statement on constipation. Gastroenterology. 2013; 144:211–217. [PubMed: 23261064]
12. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. Gastroenterology. 2008; 135:1383–1391. e1–5. [PubMed: 18789939]
13. Katz PO, Gerson LB, Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. Am. J. Gastroenterol. 2013; 108:308–328. quiz 329. [PubMed: 23419381]
14. Rubio-Tapia A, Hill ID, Kelly CP, et al. ACG clinical guidelines: diagnosis and management of celiac disease. Am J Gastroenterol. 2013; 108:656–676. quiz 677. [PubMed: 23609613]
15. Talley NJ, Vakil N. Guidelines for the management of dyspepsia. Am. J. Gastroenterol. 2005; 100:2324–2337. [PubMed: 16181387]

16. Patel P, Bercik P, Morgan DG, et al. Prevalence of organic disease at colonoscopy in patients with symptoms compatible with irritable bowel syndrome: cross-sectional survey. Scand. J. Gastroenterol. 2015; 50:816–823. [PubMed: 25636675]
17. Spiegel BM, Hays RD, Bolus R, et al. Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) gastrointestinal symptom scales. Am. J. Gastroenterol. 2014; 109:1804–1814. [PubMed: 25199473]
18. Borchard A, Schwappach DL, Barbir A, et al. A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. Ann. Surg. 2012; 256:925–933. [PubMed: 22968074]
19. Thomassen O, Storesund A, Softeland E, et al. The effects of safety checklists in medicine: a systematic review. Acta Anaesthesiol. Scand. 2014; 58:5–18. [PubMed: 24116973]
20. Gawande, A. The Checklist Manifesto: How to Get Things Right. Metropolitan Books; New York, NY: 2009.
21. Lucas GM, Gratch J, King A, et al. It's only a computer: virtual humans increase willingness to disclose. Comput. Hum. Behav. 2014; 37:94–100.
22. Ford AC, Van Zanten SV, Rodgers CC, et al. Diagnostic utility of alarm features for colorectal cancer: systematic review and meta-analysis. Gut. 2008; 57:1545–1553. [PubMed: 18676420]
23. van Zanten SV. Review alarm features in dyspepsia do not have high diagnostic accuracy for predicting upper gastrointestinal cancer. Evid. Based Med. 2007; 12:55. [PubMed: 17400647]
24. Bowrey DJ, Griffin SM, Wayman J, et al. Use of alarm symptoms to select dyspeptics for endoscopy causes patients with curable esophagogastric cancer to be overlooked. Surg. Endosc. 2006; 20:1725–1728. [PubMed: 17024539]

Summary table

What is already known:

- It is important for clinicians to inquire about “red flags” as it may identify those at risk for organic disease and who require additional diagnostic workup.
- Little is known about the effectiveness of gastroenterologists at documenting alarm features within an electronic health record-integrated practice.

What this study adds to our knowledge:

- Physicians documented less than one-third of red flags reported by patients through a computer algorithm.
- These data indicate that physicians may under report alarm features and that computerized “checklists” could complement standard medical histories to bolster clinical care.

Ms. Smith's goal for this visit is: "Figure out what is causing my abdominal pain."

HPI: Ms. Smith is a 49-year-old female who reports a history of GERD and now presents with abdominal pain. The pain first started 8 months ago, and typically lasts for 5 hours at a time. Over the past week, the pain occurred more than once a day. She describes the pain as cramping, says it is located in the left lower abdomen, and reports the pain has been "quite severe" and "very bothersome" in the past week. It does not radiate. It is associated with stress or anxiety. It usually comes on gradually. It gets better with bowel movements. The pain is somewhat relieved by reducing stress. The pain does not awaken her from sleep. She does not report early satiety. She does not report irritable bowel syndrome. She does not take aspirin or NSAIDs.

The patient does not report dysphagia, bowel incontinence, heartburn, bloating, diarrhea, constipation, nausea, or vomiting.

She does not report black stools, vomiting blood, diminished appetite, or fevers. She reports bloody bowel movements for the past 3 weeks. She has blood in her stool most of the time and has about half a cup of blood with her bowel movements. She has an unintended weight loss of 20 pounds over the past 6 months. She has no history of abdominal surgeries. There is no family history of gastrointestinal cancers.

Fig. 1.

Computer-generated history of present illness (HPI) for a fictionalized patient.

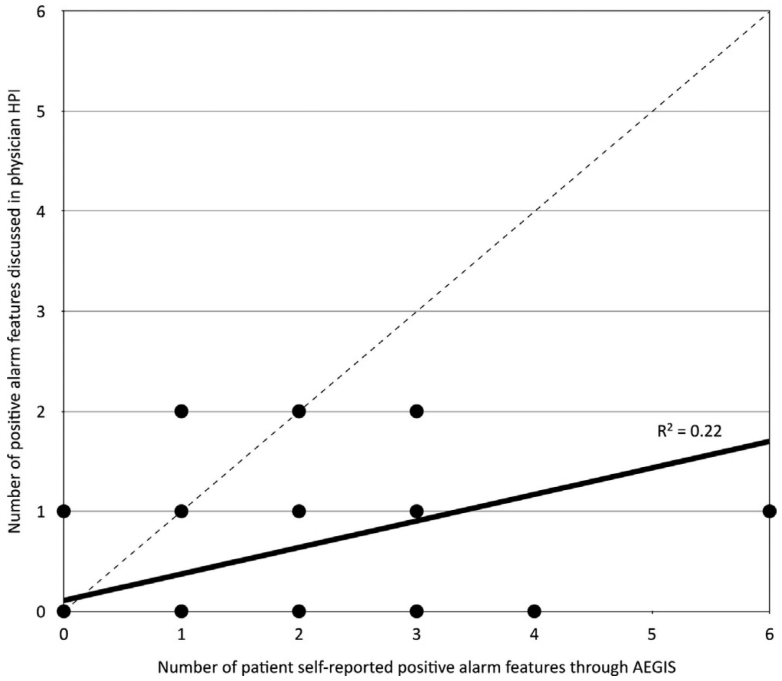


Fig. 2.

Scatterplot of number of positive alarm features mentioned in the physician history of present illness (HPI) versus that reported through Automated Evaluation of Gastrointestinal Symptoms (AEGIS). The dashed line indicates concordance between HPIs while the solid line is the ordinary least square result through the scatterplot.

Table 1

Gastrointestinal alarm features collected by AEGIS.

Gastrointestinal alarm features	Attributes
Hematochezia	Onset, frequency, amount
Melena	Onset
Hematemesis	Onset, frequency, amount
Unintentional weight loss	Onset, amount
Decreased appetite	Onset
Fevers	Onset

AEGIS, Automated Evaluation of Gastrointestinal Symptoms.

Table 2

Patient demographics.

Variable	Value (*N* = 75)
Age (years)	57.3 (13.3)
Male	50 (67)
Race/ethnicity:	
African-American	15 (20)
Asian	4 (5)
Caucasian	36 (48)
Latino	9 (12)
Native American/Hawaiian	2 (3)
Other/unknown	9 (12)
Site of care:	
University-based academic health system	28 (37)
Veterans Affairs medical center	47 (63)
Clinic visit type:	
Initial visit	39 (52)
Follow-up visit	36 (48)
Primary gastrointestinal symptom: [a,b]	
Abdominal pain	19 (28)
Bloat/gas	10 (15)
Diarrhea	6 (9)
Constipation	6 (9)
Incontinence/soilage	2 (3)
Heartburn/reflux	10 (15)
Dysphagia	12 (18)
Nausea/vomiting	2 (3)
Positive alarm features: [a,c]	
Hematochezia	9 (12)
Melena	10 (13)
Hematemesis	4 (5)
Unintentional weight loss	22 (29)
Decreased appetite	24 (32)
Fevers	2 (3)
None	35 (47)

Data are presented as mean (standard deviation) or n (%).

AEGIS, Automated Evaluation of Gastrointestinal Symptoms.

[a] Self-reported by patient through AEGIS.

[b] Eight patients reported multiple symptoms and were unable to choose the most troublesome symptom.

[c] Patients may have reported more than one positive alarm feature.

Table 3

Number of alarm features documented in the HPIs.

	Num. of positive alarm features		
	Physician HPIs	**AEGIS HPIs**	***p*-value** *
All patients (*N* = 75)	0 [0–1]	1 [0–2]	<.001
Patients presenting for an initial visit (*n* = 39)	0 [0–1]	1 [0–2]	.009
Patients who completed AEGIS within 1 week of their clinic visit (*n* = 55)	0 [0–1]	1 [0–2]	<.001

Data are presented as median [interquartile range].

AEGIS, Automated Evaluation of Gastrointestinal Symptoms; HPI, history of present illness.

* Wilcoxon signed-rank test.

Table 4

Alarm feature matching assessment.[a]

Self-reported positive alarm feature[b]	Physician HPI documented patient's self-reported positive alarm feature
Hematochezia	4/9 (44%)
Melena	1/10 (10%)
Hematemesis	2/4 (50%)
Unintentional weight loss	9/22 (41%)
Decreased appetite	5/24 (21%)
Fevers	0/2 (0%)

Data are presented as *n* (%).

AEGIS, Automated Evaluation of Gastrointestinal Symptoms; HPI, history of present illness.

[a] Limited to patients who self-reported an alarm feature through AEGIS (*n* = 40).

[b] Self-reported by patient through AEGIS.