Computer-Generated Vs. Physician-Documented History of Present Illness (HPI): Results of a Blinded Comparison

Article *in* The American Journal of Gastroenterology · December 2014

DOI: 10.1038/ajg.2014.356 · Source: PubMed

CITATIONS
27

READS
1,852

20 authors, including:

William Chey
University of Michigan
642 PUBLICATIONS **25,790** CITATIONS
SEE PROFILE

Aung Kaung
Yale University
14 PUBLICATIONS **89** CITATIONS
SEE PROFILE

Mark W. Reid
Cedars-Sinai Medical Center
37 PUBLICATIONS **962** CITATIONS
SEE PROFILE

Ken Nguyen
Cedars-Sinai Medical Center
8 PUBLICATIONS **220** CITATIONS
SEE PROFILE

Computer-Generated Vs. Physician-Documented History of Present Illness (HPI): Results of a Blinded Comparison

Christopher V. Almario, MD[1,2,3,4], William Chey, MD[5], Aung Kaung, MD[1,2], Cynthia Whitman, MPH[1,4], Garth Fuller, MS[1,4], Mark Reid, PhD[2,4], Ken Nguyen, MD[1], Roger Bolus, PhD[4], Buddy Dennis, PhD[6], Rey Encarnacion, BS[6], Bibiana Martinez, MPH[1,2,4], Jennifer Talley, MSPH[1,2,4], Rushaba Modi, MD[1,2,3], Nikhil Agarwal, MD[1,2,3], Aaron Lee, MD[1], Scott Kubomoto, MD[1], Gobind Sharma, MD[1], Sally Bolus, MS[4], Lin Chang, MD[3] and Brennan M.R. Spiegel, MD, MSHS[1,2,4]

OBJECTIVES: **Healthcare delivery now mandates shorter visits with higher documentation requirements, undermining the patient–provider interaction. To improve clinic visit efficiency, we developed a patient–provider portal that systematically collects patient symptoms using a computer algorithm called Automated Evaluation of Gastrointestinal Symptoms (AEGIS). AEGIS also automatically "translates" the patient report into a full narrative history of present illness (HPI). We aimed to compare the quality of computer-generated vs. physician-documented HPIs.**

METHODS: **We performed a cross-sectional study with a paired sample design among individuals visiting outpatient adult gastrointestinal (GI) clinics for evaluation of active GI symptoms. Participants first underwent usual care and then subsequently completed AEGIS. Each individual thereby had both a physician-documented and a computer-generated HPI. Forty-eight blinded physicians assessed HPI quality across six domains using 5-point scales: (i) overall impression, (ii) thoroughness, (iii) usefulness, (iv) organization, (v) succinctness, and (vi) comprehensibility. We compared HPI scores within patient using a repeated measures model.**

RESULTS: **Seventy-five patients had both computer-generated and physician-documented HPIs. The mean overall impression score for computer-generated HPIs was higher than physician HPIs (3.68 vs. 2.80; *P*<0.001), even after adjusting for physician and visit type, location, mode of transcription, and demographics. Computer-generated HPIs were also judged more complete (3.70 vs. 2.73; *P*<0.001), more useful (3.82 vs. 3.04; *P*<0.001), better organized (3.66 vs. 2.80; *P*<0.001), more succinct (3.55 vs. 3.17; *P*<0.001), and more comprehensible (3.66 vs. 2.97; *P*<0.001).**

CONCLUSIONS: **Computer-generated HPIs were of higher overall quality, better organized, and more succinct, comprehensible, complete, and useful compared with HPIs written by physicians during usual care in GI clinics.**

Am J Gastroenterol 2015; 110:170–179; doi:10.1038/ajg.2014.356; published online 2 December 2014

INTRODUCTION

Electronic health records (EHRs) have the potential to improve outcomes and quality of care, yield cost savings, and increase engagement of patients with their own health care (1). When successfully integrated into clinical practice, EHRs automate and streamline clinician workflows, narrowing the gap between information and action that can result in delayed or inadequate care (2). In recent years, EHR adoption has proceeded at an accelerated rate, fundamentally altering the way healthcare providers document, monitor, and share information (3).

[1]Division of Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, California, USA; [2]Division of Gastroenterology, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA; [3]Division of Digestive Diseases, UCLA, Los Angeles, California, USA; [4]Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Los Angeles, California, USA; [5]Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, USA; [6]UCLA Computing Technology Research Laboratory (CTRL), Los Angeles, California, USA. **Correspondence:** Brennan M.R. Spiegel, MD, MSHS, Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Health Services Research, Cedars-Sinai Health System, Pacific Theaters Building, 116 North Robertson Boulevard, 4th Floor, Los Angeles, California 90048, USA. E-mail: brennan.spiegel@cshs.org
Received 8 August 2014; accepted 14 October 2014

Although there is evolving evidence that EHRs can modestly improve clinical outcomes (4), EHR systems were principally designed to support the transactional needs of administrators and billers, and less so to nurture the relationship between patients and their providers (5). Nowhere is this more apparent than in the ability of EHRs to handle unstructured, free-text data of the sort found in the history of present illness (HPI). As the proximal means of assessing a patient's current illness experience, the HPI proceeds as an open-ended interview, eliciting patient input and summarizing the information as free text within the patient record (6). Healthcare providers use the HPI to execute care plans and to document a foundational reference for subsequent encounters. In addition, the HPI is transformed by a clinical coding specialist into more structured data linked to payment and reimbursement (7).

In an EHR-integrated practice environment, the physician enters HPIs directly into the patient record. Studies examining EHR-based encounters reveal that physicians often engage in computer-related behaviors that patients find unsettling, including performing computational tasks at the computer screen (8–10). In addition, the quality of physician-captured HPIs remains highly variable, with some found to be inaccurate or incomplete (11).

There is room for solutions that increase the validity of the HPI, alleviate inconsistencies in free text captured at the point of care, and reduce the number of computational tasks. In this study, we compared the quality of physician HPIs composed during usual care with computer-generated HPIs created by an algorithm that "translates" patient symptoms into a narrative HPI written in language familiar to clinicians. By comparing physician-documented and computer-generated HPIs independently obtained on the same patients, we aimed to assess relevance, clarity, and completeness of the HPIs, as well as evaluate compliance with reimbursement standards achieved between the two methods.

METHODS

Study overview

We compared HPIs generated through two methods on the same patients: (i) physician-documented HPIs documented in the EHR; and (ii) computer-generated HPIs created by algorithms trained to collect and display a medical history in narrative form. Blinded physician raters evaluated the HPIs without knowledge about the purpose of the study. We conducted the study in gastrointestinal (GI) clinics at the University of California at Los Angeles (UCLA) and the West Los Angeles Veterans Administration (WLAVA) Medical Center. We selected these clinics because GI patients frequently report chronic and discrete symptoms amenable to computerized analysis (12), and to evaluate how a computer-generated HPI compares with HPIs obtained from subspecialists in diverse secondary and tertiary care clinics.

Automated Evaluation of Gastrointestinal Symptoms (AEGIS)

We tested an HPI computer algorithm called AEGIS. AEGIS is available through *My GI Health*, a publically available patient–provider portal created by our groups at Cedars-Sinai Medical Center and the University of Michigan. See www.MyGIHealth.org for more information about the portal. *My GI Health* was identified by the PCORI (Patient-Centered Outcomes Research Institute) as one of 11 model systems in a landscape review commissioned for the 2013 PCORI Patient-Reported Outcomes Infrastructure Workshop (13), and is supported in part with funding from the NIH (National Institutes of Health) Patient-Reported Outcomes Measurement Information System (PROMIS; www.nihPROMIS.org) (14).

My GI Health uses AEGIS to collect patient information regarding GI symptoms. Once on the portal, patients report which among eight GI symptoms they have experienced, including: (i) abdominal pain, (ii) bloat/gas, (iii) diarrhea, (iv) constipation, (v) bowel incontinence, (vi) heartburn/reflux, (vii) disrupted swallowing, and (viii) nausea/vomiting. These symptoms are based on the NIH PROMIS framework of GI symptoms that we previously developed using data from over 2,000 subjects (12,14,15). If a patient reports multiple symptoms in AEGIS, then the system prompts the user to select the most bothersome symptom, when possible.

For each reported GI symptom, AEGIS guides patients through GI PROMIS computerized adaptive testing questionnaires to measure symptom severity. The PROMIS scores are converted to a symptom "heat map" (**Figure 1**) that visually compares a patient's symptoms with the general US population (12,14).

AEGIS also guides patients through questions drawn from a library of over 300 symptom attributes measuring the timing, severity, frequency, location, quality, bother, and character of their GI symptoms, along with relevant comorbidities, family history, and alarm features. We selected these clinical factors based on ideal HPIs presented in standard medical texts (16), and further shaped the computer-generated narratives with input from experienced gastroenterologists (W.C., L.C., and B.M.R.S.). To tailor the questions to each respondent's unique symptom experience, patients only answer a subset of questions conditional upon their answers. For this study, we did not time how long it took patients to complete AEGIS. However, in our previous pilot testing of AEGIS, most patients took ~10–15 min to complete the HPI questionnaire based on time-stamped start and stop times.

Once the questions are completed, the information is transformed into a full narrative HPI (see **Figure 2** and **Appendix** for examples). For patients who report only one active symptom, AEGIS structures the HPI so that the first paragraph discusses the chief complaint, the second comments on pertinent negative symptoms, and the third describes pertinent alarm features and relevant family history.

For patients with multiple active symptoms, AEGIS discusses the patient's most bothersome symptom in the first paragraph, and subsequent paragraphs are arranged in order of PROMIS symptom severity scores. In this manner, AEGIS prioritizes the information using a combination of patient gestalt and empirical data.

Setting

We performed a cross-sectional study using a paired sample design. We enrolled patients who visited GI clinics at WLAVA or UCLA. The WLAVA GI clinic is an academic teaching practice staffed by GI attending physicians. GI subspecialty fellows and internal medicine residents conduct initial evaluations in most

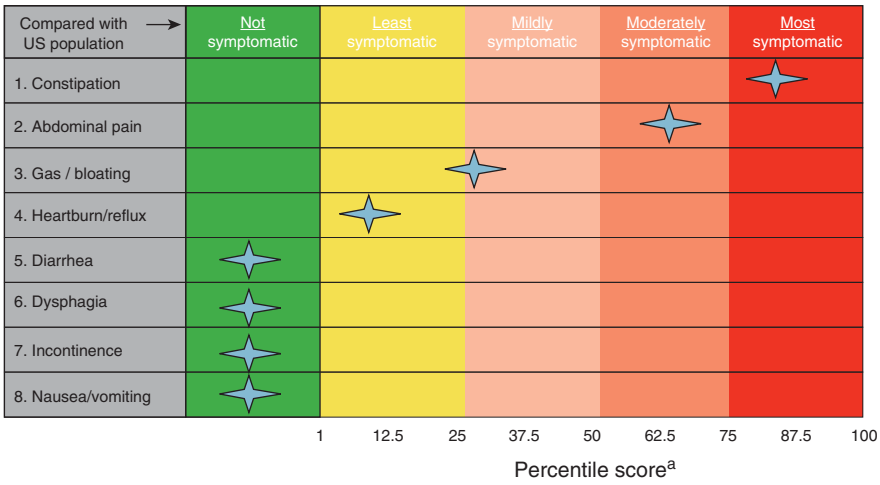

Figure 1. Sample "heat map" patient report of GI PROMIS scores. Patients complete PROMIS items on *My GI Health*, and the results are converted into a symptom score visualization. Patients' scores are compared with the general US population with benchmarks to add interpretability to the scores, similar to a lab test. GI, gastrointestinal; PROMIS, Patient-Reported Outcomes Measurement Information System.

History of the presenting illness [-]

HPI: Mr. Smith is a 59-year-old male with a history of non-insulin-dependent diabetes mellitus who now presents with diarrhea. He first began having diarrhea 5 months ago. The diarrhea has gradually worsened over time. He has loose stools 6–7 days per week. On the days he has diarrhea, he has 5–7 loose or watery stools per day. His bowel movements range from "fluffy pieces with ragged edges, a mushy stool" (Bristol Type 6) to "watery, no solid pieces" (Bristol Type 7). His most common stool type is Bristol Type 7 (watery, no solid pieces). He often has urgency with bowel movements. He describes stool as foul smelling, and greasy or oily. He is not taking any medications associated with diarrhea. He does not report a history of irritable bowel syndrome, thyroid disease, or gastrointestinal cancer. He has not yet tried any therapy for diarrhea.

He also reports bowel incontinence. The bowel incontinence first started 3 months ago and occurs once a week. His usual bowel habit is 7 liquid bowel movements per day. He is incontinent for a small amount of stool. When this occurs, he usually feels an urge to defecate before the incontinence, but cannot make it to the bathroom in time. He sometimes thinks he is going to pass gas but passes stool or liquid instead. He wears a pad rarely. These symptoms interfere with his day-to-day activities somewhat. He does not have urinary incontinence. He does not have a history of anal sphincter injury or surgery. He has not yet tried any therapy for bowel incontinence. He has not yet tried any dietary interventions for bowel incontinence.

He does not report abdominal pain, constipation, or dysphagia. He also has not recently had nausea, vomiting, bloating, abdominal swelling, heartburn, or regurgitation. He does not report a history of neurological conditions like stroke, Parkinson's disease, myasthenia gravis, or multiple sclerosis.

He does not report blood in his bowel movements, black stools, vomiting blood, unintended weight loss, diminished appetite, or fevers. He has no history of abdominal surgeries. There is a family history of colorectal cancer.

Figure 2. Sample of a computer-generated history of present illness (HPI).

patients at WLAVA, and generate HPIs by typing information into the VA Computerized Patient Record System. In contrast, attending physicians primarily staffed the UCLA GI clinics, including practices at Ronald Reagan UCLA Medical Center and Santa Monica-UCLA Medical Center. Physicians in the UCLA clinics either hand typed or dictated HPIs into the EHR developed by Epic Systems Corporation (Verona, WI).

Patients

We enrolled patients aged ≥18 years who were scheduled for an initial or follow-up visit. Patients were required to have one or more symptoms at the time of their visit. We excluded asymptomatic patients, or those seeking consultation solely for abnormal blood tests, cancer screening, or other nonsymptomatic indications. Patients were required to read and write English and possess basic computing skills.

All patients first received usual care in their respective clinics before completing AEGIS. Although computer-generated HPIs would be more useful before a visit in real practice, we could not administer AEGIS before the visit because it would 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. They were either recruited directly in clinic or through mailed recruitment materials. During clinic, patients who met inclusion criteria were invited to use *My GI Health* on a computer after completion of their physician visit. For the remaining patients not recruited in clinic, their EHR charts were reviewed the following day, and those meeting inclusion criteria received a letter inviting them to visit *My GI Health* online. Patients from UCLA were recruited between January and March 2014 through mailed recruitments.

Primary outcomes

We assessed both the computer-generated and physician-documented HPIs using two techniques: blinded assessment by physicians using metrics based on the Physician Documentation Quality Instrument (17) and blinded evaluation by a billing compliance officer.

Physician evaluations. We recruited a panel of 48 reviewers from across the United States and Canada that included 16 board-certified gastroenterologists, 16 board-certified internists, and 16 internal medicine residents. Reviewers were not informed about AEGIS or that some HPIs were computer generated. Reviewers were only instructed that we were auditing the quality of HPIs generated from patients in GI clinics.

Each reviewer received a set of 10 randomly selected HPIs from 10 unique patients (5 each for computer and physician HPIs). After stripping all identifying information from the HPIs, we distributed test sets online using SurveyMonkey (SurveyMonkey, Palo Alto, CA). Reviewers employed a 5-point Likert scale (1=extremely poor; 5=excellent) to measure HPIs across six domains: (i) overall impression, (ii) thoroughness, (iii) usefulness, (iv) organization, (v) succinctness, and (vi) comprehensibility (17). Three blinded reviewers, including one gastroenterologist, one internist, and one internal medicine resident, independently evaluated each HPI; this ensured that multiple reviewers with different background and training reviewed each HPI.

CMS level of complexity. The level of complexity of HPIs partly determines Centers for Medicare & Medicaid Services (CMS) reimbursement rates for a patient encounter. CMS suggests that HPIs document a range of symptom-based elements, including: (i) location (not applicable for diarrhea, constipation, incontinence, or nausea/vomiting in this study), (ii) quality, (iii) severity, (iv) duration, (v) timing, (vi) context, (vi) modifying factors, and (viii) associated signs and symptoms (18). An experienced CMS compliance auditor blindly assessed all HPIs in random order to enumerate which elements were documented for the primary GI symptom. The compliance auditor was not informed that some HPIs were computer generated.

Secondary outcomes

We compared whether the chief symptom complaint documented in the physician HPI matched the patient's self-reported most bothersome symptom reported to the computer. In order to ensure an unbiased determination of the chief complaint, the blinded compliance auditor independently reported the chief complaint for each HPI.

We also compared the total number of active GI symptoms addressed in the computer with physician HPIs. A single blinded physician reviewer not associated with the study team enumerated the active GI symptoms described in each HPI. Given that a patient's symptoms may change over time, we limited these secondary analyses to patients who completed AEGIS within 1 week of their GI clinic visit.

Sample size calculation

Our primary outcome measure was the overall impression score of the HPI as determined by blinded physicians. No data exist regarding the minimally important difference on the 5-point global impression scale. Therefore, we calculated a minimum sample size required to achieve an effect size of 0.5 s.d. (19). Assuming a two-tailed 5% significance level with a power of 80%, the minimum sample size needed to show an effect size of 0.5 was 32 total patients.

Statistical analysis

We used REDCap (Research Electronic Data Capture) hosted at UCLA to collect study data (20). Statistical analyses were performed using Stata 13.1 (StataCorp LP, College Station, TX). A two-tailed P value of <0.05 was considered significant in all analyses. For bivariate analyses, we used either the two-sample t-test or analysis of variance test to compare means between groups.

The primary outcome was difference in mean overall impression score between computer-generated and physician HPIs. We used repeated measures analysis of covariance to generate an adjusted P value and to evaluate differences between groups while adjusting for potential confounding factors, including 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.

We used repeated measures analysis of covariance to compare the number of CMS-recommended HPI elements present in the computer with physician HPIs. To evaluate for differences in the proportion of HPIs with ≥4 elements (considered an "extended HPI" by CMS and important for reimbursement purposes), we used the two-sample test of proportions. For the analysis comparing the number of active GI symptoms between HPI types, we generated a scatterplot and calculated R^2 of the ordinary least squares regression line. The UCLA (IRB 13-000337) and WLAVA (IRB PCC 2014-020138) Institutional Review Boards approved this study. We employed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines to report this study (21).

RESULTS

Patients

Table 1 presents information regarding the 75 study patients, each of whom contributed two HPIs for a total of 150 HPIs. Of these,

Table 1. Patient demographics

Variable	Value (of *N*=75 total)
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)
Physician HPI author	
GI attending	26 (35)
GI fellow	36 (48)
Internal medicine resident	10 (13)
GI physician assistant	3 (4)
Physician HPI input method	
Verbally dictated	6 (8)
Typed	69 (92)
Clinic visit type	
Initial visit	39 (52)
Follow-up visit	36 (48)
Site of care:	
University-based academic health system	28 (37)
Veterans Affairs medical center	47 (63)
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)
Swallowing	12 (18)
Nausea/vomiting	2 (3)

GI, gastrointestinal; HPI, history of present illness.
Data are presented as mean (s.d.) or *n* (%).
[a]Self-reported by patient as most troublesome symptom through Automated Evaluation of Gastrointestinal Symptoms (AEGIS).
[b]Eight patients reported multiple symptoms and were unable to choose the most troublesome symptom.

29 patients were directly recruited in the WLAVA GI clinic and completed AEGIS before leaving. The remaining 46 patients were enrolled through mailed recruitment materials following their visit. Among patients recruited by mail, 8% (18/220) of WLAVA patients and 7% (28/428) of UCLA patients completed AEGIS a median of 7 days after their visit (range: 2–29 days). Among all recruited patients, 28% (21/75) reported only one active symptom through AEGIS; the remainder noted multiple symptoms.

Physicians

The 75 physician HPIs were written by 6 GI attending physicians, 9 GI subspecialty fellows, 7 internal medicine residents, and 1 GI physician assistant—a total of 23 unique providers. Of the physician HPIs, 92% (69/75) were typed directly into the EHR, whereas the remaining 8% (6/75) were verbally dictated. No differences were seen between the typed HPIs and the dictated HPIs with respect to scores for overall impression (P=0.38), completeness (P=0.47), relevance (P=0.79), organization (P=0.06), succinctness (P=0.16), or comprehensibility (P=0.06). When comparing the 28 physician HPIs written at UCLA with the 47 physician HPIs composed at the WLAVA, no differences were seen in overall impression (P=0.17), completeness (P=0.26), relevance (P=0.46), organization (P=0.98), succinctness (P=0.53), or comprehensibility (P=0.98).

When comparing HPIs written by GI attendings, GI fellows, and residents/physician assistants, no differences were noted in completeness (P=0.10), relevance (P=0.14), organization (P=0.29), succinctness (P=0.67), or comprehensibility (P=0.41). However, attending physicians generated HPIs that had significantly lower overall impression scores (mean 2.50, s.d. 0.77) than those generated by fellows (mean 2.90, s.d. 0.76) and residents/physician assistants (mean 3.10, s.d. 0.44; P=0.03). Thus, in primary and subgroup multivariate analyses, we controlled for differences observed in HPI quality indices across HPI author groups. We did not control for individual differences in physician performance because HPI quality indices were homoscedastic across groups in most cases, and no subgroups demonstrated statistically significant deviations from normality.

Primary analysis results

Table 2 presents the HPI assessments comparing the computer HPIs with physician HPIs. Compared with the physician HPIs, blinded raters judged the computer HPIs to have significantly higher scores for overall impression (repeated measures analysis of covariance adjusted P<0.001), completeness (adjusted P<0.001), relevance (adjusted P<0.001), organization (adjusted P<0.001), succinctness (adjusted P<0.001), and comprehensibility (adjusted P<0.001). The blinded compliance auditor identified documentation of a clear chief complaint in 100% (75/75) of computer HPIs vs. 75% (56/75) of physician HPIs (P<0.001). The auditor found documentation of more CMS-recommended symptom criteria in computer than in physician HPIs with respect to the chief complaint (adjusted P<0.001; **Table 2**). More computer than physician HPIs met CMS criteria for extended status (≥4 elements; 100% (75/75) vs. 88% (49/56); P=0.002). This difference lost significance when focusing only on initial visits (100% (39/39) vs. 94% (30/32); P=0.11).

Subgroup analyses

In subgroup analysis focused only on initial visits (n=39), AEGIS HPIs were judged to be of higher overall quality (adjusted P<0.001), more complete (adjusted P<0.001), more relevant

Table 2. HPI ratings

	Physician HPIs (*N*=75)	Computer-generated HPIs (*N*=75)	Adjusted *P* value[a]
Overall impression[b]	2.80 (0.75)	3.68 (0.61)	<0.001
Completeness[b]	2.73 (0.75)	3.70 (0.59)	<0.001
Relevance[b]	3.04 (0.68)	3.82 (0.54)	<0.001
Organization[b]	2.80 (0.80)	3.66 (0.63)	<0.001
Succinctness[b]	3.17 (0.60)	3.55 (0.69)	<0.001
Comprehensibility[b]	2.97 (0.79)	3.66 (0.66)	<0.001
Number of CMS-recommended elements present in HPI[c,d]	5.27 (1.52)	6.05 (0.98)	<0.001

CMS, Centers for Medicare & Medicaid Services; HPI, history of present illness.
Data are presented as mean (s.d.).
[a]Adjusted using a repeated measures model. Variables in the model included patient age, sex, author of physician HPI, physician HPI input method, clinic visit type, and site of care.
[b]Assessed by blinded physician reviewers.
[c]Assessed by a blinded physician compliance auditor.
[d]Nineteen patients were excluded given their physician HPI had an unclear primary gastrointestinal (GI) symptom.

Table 3. Physician HPI's primary GI symptom matched patient's self-reported primary GI symptom[a]

Self-reported primary GI symptom[b]	Physician HPI's primary GI symptom matched patient's self-reported primary GI symptom	Any comment in physician's HPI about patient's self-reported primary GI symptom
Abdominal pain	10/14 (71)	13/14 (93)
Bloat/gas	4/7 (57)	5/7 (71)
Diarrhea	1/5 (20)	5/5 (100)
Constipation	2/4 (50)	3/4 (75)
Incontinence/soilage	0/2 (0)	2/2 (100)
Heartburn/reflux	3/6 (50)	6/6 (100)
Swallowing	8/9 (89)	9/9 (100)
Nausea/vomiting	0/2 (0)	1/2 (50)

GI, gastrointestinal; HPI, history of present illness.
Data are presented as *n* (%).
[a]Limited to patients who completed Automated Evaluation of Gastrointestinal Symptoms (AEGIS) within 7 days of their GI clinic visit and who reported a most troublesome symptom through AEGIS (*n*=49).
[b]Self-reported by patient as most troublesome GI symptom through AEGIS.

(adjusted *P*<0.001), better organized (adjusted *P*<0.001), more succinct (adjusted *P*=0.002), and more comprehensible (adjusted *P*<0.001). Computer HPIs maintained higher scores for overall impression (adjusted *P*<0.001), completeness (adjusted *P*<0.001), relevance (adjusted *P*<0.001), organization (adjusted *P*<0.001), and comprehensibility (adjusted *P*<0.001) when excluding reviews from GI physicians (who might rate GI HPIs differently from non-GI specialists). There was a marginal effect with succinctness (adjusted *P*=0.08).

Secondary analyses

We evaluated whether the chief symptom complaint documented in the physician HPI matched the self-reported chief complaint in AEGIS. Because symptoms can vary over time, we limited this analysis to patients who completed AEGIS within 7 days of their GI clinic visit and reported a most bothersome symptom (*n*=49). The physician-documented chief complaint matched the patient's self-report in 57% (28/49). Of the physician HPIs, 90% (44/49) made at least some mention of the patient's self-reported chief complaint. **Table 3** lists these analyses according to GI symptom.

Finally, we compared the number of GI symptoms discussed in the physician HPI with the number of active symptoms self-reported by the patient through AEGIS. When compared side by side, computer HPIs documented more active symptoms than physician HPIs (**Table 4** and **Figure 3**).

Table 4. Comparing number of active GI symptoms addressed in the physician HPI with the number of active GI symptoms reported by patients through AEGIS[a]

No. of active GI symptoms reported through AEGIS	No. of active GI symptoms addressed in physician HPI[b]
1 (*n*=16)	1.6 (0.7)
2 (*n*=12)	2.0 (1.1)
3 (*n*=10)	2.2 (1.0)
4 (*n*=5)	1.8 (0.8)
5 (*n*=7)	2.9 (1.3)
6 (*n*=4)	2.5 (1.3)
7 (*n*=1)	5.0 (0)
8 (*n*=0)	—

AEGIS, Automated Evaluation of Gastrointestinal Symptoms; GI, gastrointestinal; HPI, history of present illness.
Data are presented as mean (s.d.).
[a]Limited to patients who completed AEGIS within 7 days of their GI clinic visit (*n*=55).
[b]Determined by a single blinded physician.

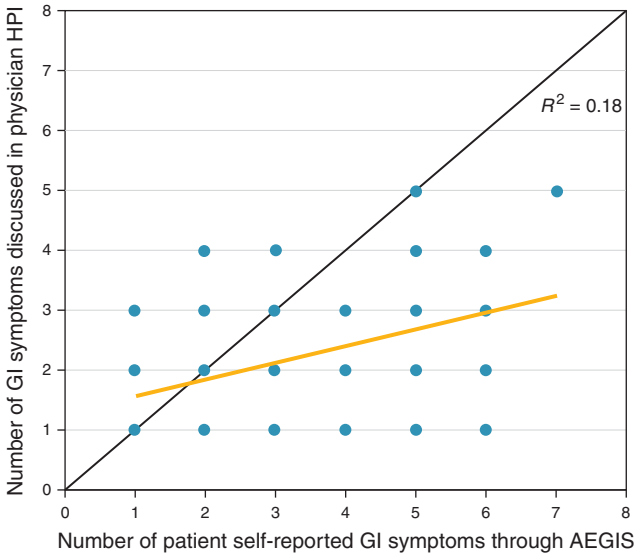

Figure 3. Scatterplot of number of active gastrointestinal (GI) symptoms addressed in the physician history of present illness (HPI) vs. number of active GI symptoms reported through Automated Evaluation of Gastrointestinal Symptoms (AEGIS). The diagonal line indicates concordance between HPIs. The yellow line is the ordinary least square result through the scatterplot, indicating a lower than expected correlation, with AEGIS capturing more active symptoms, on average, than physician HPIs.

DISCUSSION

We studied whether a computer could gather and present patient-entered data as an HPI narrative, and compared the resulting HPIs with those generated by physicians during usual care. We found that blinded raters deemed the computer-generated HPIs to be of higher quality, more comprehensive, better organized, and with greater relevance compared with physician-documented HPIs. These results offer initial proof of principle that a computer can create meaningful and clinically relevant HPIs, and the findings are consistent with recent evidence that patients are comfortable disclosing health information to "virtual human" interviewers as supportive and "safe" interaction partners (22).

In contrast, these results do not indicate that computers could ever replace healthcare providers. The art of medicine requires that physicians connect with and empower patients, interpret complex and oftentimes confusing data, render diagnoses despite imperfect information, and communicate in an effective manner. Nonetheless, this study suggests that computers can at least lift some burden by collecting and organizing data to help clinicians focus on what they do best—practicing the distinctly human art of medicine.

Healthcare providers typically summarize HPIs as free text within the EHR, often during the patient visit (10). Data reveal that clinicians spend up to one-third of the patient encounter looking at a computer and, as a result, miss important nonverbal cues from the patient (8,9). Studies also show that the quality of physician-captured HPIs remains highly variable, with many HPIs proving inaccurate or incomplete (11). The AEGIS algorithms tested in this study offers one model to begin addressing these issues, as they create HPIs informed by patient self-report and PROMIS measures, thereby increasing the validity of the HPI and reducing data inconsistencies. Moreover, use of computer-generated HPIs may improve the patient–physician relationship by informing the clinician in advance of a patient encounter and freeing more time by reducing computational tasks.

Beyond the primary goal of collecting and organizing facts about a patient's symptom experience, HPIs are also used to justify the level of complexity billed by medical providers (18,23). In this study, we found that every computer-generated HPI met criteria for the highest level of complexity (18), as determined by a CMS compliance auditor. In contrast, 88% of physician HPIs met the CMS criteria. This suggests that computer-generated HPIs achieve documentation requirements, but we must also recognize that just because the computer collects more information does not mean it is more accurate—this is the subject of future research.

We discovered a frequent mismatch between the patient's self-reported most bothersome symptom and the physician-documented chief complaint. Namely, only 57% of patients had a physician HPI that matched their self-reported chief complaint. It is possible that some patients did not mention their most bothersome symptom to their physicians, or perhaps felt more comfortable reporting it to the computer. However, it is also important to emphasize that patients in this study completed AEGIS after their clinic visit. It is possible that the chief complaint at the time of the clinic visit was adequately addressed and treated by the time they completed AEGIS. Moreover, symptoms may have changed between the clinic visit and the time the patient completed AEGIS. To address these issues, we limited this secondary analysis to those who completed AEGIS within 1 week of their clinic appointment. Although we cannot know from this study why there is a mismatch, computer HPIs generated before visit and available to the physician

during the visit might improve upon this mismatch and enhance patient-centered care, but that must be formally investigated.

This study has important limitations. First, just because an HPI receives high ratings by a blinded reviewer does not mean it is accurate. The reviewers in this study did not have access to the patients themselves, their medical chart, or recordings from the patient encounter—only to the HPIs. Nonetheless, the computer-generated HPIs were entirely based on patient self-report, without intervening interpretation by anyone else. The computer HPI is only meant as a starting point for the patient dialog, and can be edited by the physician. Second, although we selected reviewers who were not aware of our study, it is possible that some reviewers might have detected a pattern indicating structural similarities between the computer HPIs. To counter this possibility, we only assigned 10 HPIs to each reviewer, minimizing the ability to identify a pattern among the small set of randomly presented HPIs. Third, patients completed the computer HPI after—not before—their visit. Administering AEGIS before the visit was not feasible as this would prime the patient and potentially unblind the study. It is possible that the physician visit might have affected how patients subsequently completed AEGIS, but this would have little effect on the overall structure, narrative, and appearance of the HPI itself—the attributes judged by the blinded reviewers. Fourth, only 8% of the physician HPIs were inputted into the EHR through verbal dictation. Although the verbally dictated HPIs had higher scores from blinded physician raters compared with typed HPIs, it was not statistically significant as our study was underpowered with respect to this outcome. It is certainly possible that dictated HPIs could have compared favorably with AEGIS HPIs. Fifth, we did not measure patient or provider satisfaction with AEGIS—only the quality of the resulting HPIs. We are currently performing a longitudinal controlled trial of AEGIS vs. usual care, and are collecting patient and provider satisfaction data with this alternative workflow. Future reports will prove more information on these important outcomes. Finally, only a minority of invited patients completed the computer HPI, raising the possibility of a selection bias. In particular, nonrespondents might have low health or computer literacy. However, it is unclear whether nonresponders with adequate literacy would generate systematically different computerized histories vs. responders with adequate literacy.

In summary, this study demonstrates how computer algorithms can create HPIs that provide clinicians with meaningful, relevant clinical information while remaining easy to understand and succinct. Future research will test whether computer-generated HPIs can enhance the physician–patient relationship, increase patient satisfaction, lead to more patient-centered care, or increase patient engagement in their own health care. Aside from affecting clinical care, this structured data capture may have implications for research and population health by creating more easily audited HPIs, and may also support emerging care models such as telehealth or other electronic consultations.

CONFLICT OF INTEREST

Guarantor of the article: Brennan M.R. Spiegel, MD, MSHS.

Specific author contributions: Christopher V. Almario and Aung Kaung: planning and conducting the study, collecting and interpreting data, drafting the manuscript, and approval of final draft submitted; William Chey, Cynthia Whitman, and Brennan M.R. Spiegel: planning and conducting the study, interpreting data, drafting the manuscript, and approval of final draft submitted; Garth Fuller and Mark Reid: interpreting data, drafting the manuscript, and approval of final draft submitted; Ken Nguyen, Buddy Dennis, Rey Encarnacion, Bibiana Martinez, Jennifer Talley, Rushaba Modi, Nikhil Agarwal, Aaron Lee, Scott Kubomoto, Gobind Sharma, and Sally Bolus: planning and conducting the study and approval of final draft submitted; Roger Bolus and Lin Chang: planning and conducting the study, interpreting data, and approval of final draft submitted.

Financial support: This study was supported by 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 T32DK07180-39) 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.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Electronic health record adoption has accelerated, thereby fundamentally altering the way healthcare providers document, monitor, and share information.
- ✓ The quality of the physician-captured history of present illness (HPI) is highly variable, and at times inaccurate or incomplete.

WHAT IS NEW HERE

- ✓ Automated Evaluation of Gastrointestinal Symptoms (AEGIS) offers initial proof of principle that a computer algorithm can systematically collect patient information and create HPIs that provide clinicians with meaningful, relevant clinical information.

REFERENCES

1. Blumenthal D. Launching HITECH. N Engl J Med 2010;362:382–5.
2. 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.
3. 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;32:1478–85.
4. Goldzweig CL, Towfigh A, Maglione M *et al.* Costs and benefits of health information technology: new trends from the literature. Health Aff (Millwood) 2009;28:w282–w293.
5. van der Lei J, Moorman PW, Musen MA. Electronic patient records in medical practice: a multidisciplinary endeavor. Methods Inf Med 1999;38:287–8.
6. Sapira JD. The Art and Science of Bedside Diagnosis Urban & Schwarzenberg: Baltimore, 1990.
7. The MITRE Corporation. Electronic Health Records Overview 2006.
8. Asan O, DS P, Montague E. More screen time, less face time - implications for EHR design. J Eval Clin Pract 2014;e1–6.
9. Montague E, Asan O. Dynamic modeling of patient and physician eye gaze to understand the effects of electronic health records on doctor-patient communication and attention. Int J Med Inform 2014;83:225–34.

10. Ventres W, Kooienga S, Vuckovic N *et al.* Physicians, patients, and the electronic health record: an ethnographic analysis. Ann Fam Med 2006;4:124–31.
11. Redelmeier DA, Schull MJ, Hux JE *et al.* Problems for clinical judgement: 1. Eliciting an insightful history of present illness. CMAJ 2001;164:647–51.
12. Spiegel BM. Patient-reported outcomes in gastroenterology: clinical and research applications. J Neurogastroenterol Motil 2013;19:137–48.
13. PCORI. Patient Reported Outcome Infrastructure Workshop. Atlanta, Georgia 2013.
14. 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;109:1804–14.
15. Nagaraja V, Hays RD, Khanna PP *et al.* Construct validity of the Patient Reported Outcomes Measurement Information System (PROMIS®) gastrointestinal symptom scales in systemic sclerosis. Arthritis Care Res (Hoboken) 2014;66:1725–30.
16. Bickley L, Szilagyi PG. Bates' Guide to Physical Examination and History-Taking Lippincott Williams & Wilkins: 2012.
17. Stetson PD, Bakken S, Wrenn JO *et al.* Assessing Electronic Note Quality Using the Physician Documentation Quality Instrument (PDQI-9). Appl Clin Inform 2012;3:164–74.
18. Medicare Learning Network. 1995 Documentation Guidelines for Evaluation and Management Services 1995.
19. Cohen J. Statistical Power Analysis for the Behavioral Science Academic Press: London, 1960.
20. Harris PA, Taylor R, Thielke R *et al.* Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.
21. von Elm E, Altman DG, Egger M *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61:344–9.
22. Gale M, Lucas A, Gratch J *et al.* It's only a computer: virtual humans increase willingness to disclose. Comp Hum Behav 2014;37:94–100.
23. Center for Medicare & Medicaid Services. CPT/HCPCS codes included in range 99201-99205 2014.

APPENDIX

Sample physician-documented vs. computer-generated history of present illnesses (HPIs) on same patients using the Automated Evaluation of Gastrointestinal Symptoms (AEGIS) algorithms. First example from physician HPI judged in upper third of quality, followed by examples in middle and lower third, respectively.

(A) Physician-generated HPI (overall impression rating in the upper third) vs. AEGIS HPI.

Physician HPI

A 67-year-old male with a past medical history significant for hyperlipidemia who presents for evaluation of dysphagia. The patient is otherwise healthy. He notes ~3 years ago having a period where he was having dysphagia. This started after the patient had a piece of chicken at a barbeque and felt it became stuck. For the next 3 days he states he was having difficulties with swallowing, and was spitting up significant amounts of frothy liquid. He did not seek care at the emergency room, but did seek the care of his gastroenterologist, who proceeded to perform an endoscopy in 2010. The patient is unable to recall what was seen or if any additional interventions were performed during the endoscopy, but noted resolution of his symptoms after the endoscopy. He noted that for the next 1–2 years he was symptom free and was able to eat normally. Then, in the past year he would have occasional episodes where he would feel food would occasionally get caught after swallowing. Typically most episodes resolved on their own and he would then be able to eat normally.

However, late last year, he had an acute attack of his dysphagia, and one evening noted feeling like he choked on his dinner of asparagus. He was brought to his local emergency room, but later left after 5 h without being seen. He did see his gastroenterologist who then performed a upper endoscopy and performed a wire-guided Savary dilation up to 17 mm. It was noted on the endoscopy report that a somewhat tortuous esophagus was seen but otherwise no strictures were noted. He did not note any significant improvement in his symptoms after the dilation was performed. He also underwent esophagrams and motility studies that were only noted for poor esophageal motility with a solid food bolus. He had a repeat esophagogastroduodenoscopy (EGD) where esophageal biopsies were performed and showed reflux esophagitis. Early this year, he started developing worsening chest pain, described as a tight sensation in the mid chest that would radiate bilaterally upwards. He again went to an emergency room (ER) for the chest pain and was ruled out for cardiac causes. He also noted the onset of sharp right upper quadrant (RUQ) abdominal pain at the same time and had a gallbladder ultrasound performed in the ER that was normal. He was then referred here and underwent esophageal manometry with findings of abnormal relaxation of the lower esophageal sphincter consistent with outlet obstruction. Given these findings, he had a repeat EGD performed a week later with 25 units of botox injected at the gastroesophageal (GE) junction in four quadrants. Since the botox injection, the patient notes no improvement in his dysphagia. He did note improvement in his chest pain after botox injection. Currently, he maintains a mostly liquid diet, with some cream of wheat, small bites of crackers, and juices/protein shakes, and Ensure supplements. He tried a small bite of fish several days ago and felt it was stuck and would not pass. He notes with liquids that occasionally he would feel it building up but will eventually pass through. The patient has noted over 15 lbs of weight loss in the past 3 months.

AEGIS HPI

A 67-year-old male who reports no relevant gastrointestinal (Gl) conditions and now presents with dysphagia. He first noticed feeling dysphagia 3 months ago when he initially had difficulty swallowing solid foods. It quickly worsened over time. Now he also has trouble swallowing hard foods, but he can tolerate soft foods and liquids. He is not missing teeth. If he chews food thoroughly he can swallow but with much difficulty. He has dysphagia every day of the week and with all swallows. He reports the dysphagia has been very much bothersome. He is not sure if he has a problem initiating a swallow. He feels like the food gets stuck around his throat, sternal notch, upper sternum, mid-sternum, and lower

sternum. He rarely chokes or gags on liquid or food with swallows. He always has odynophagia with swallows.

He also reports gastroesophageal reflux disease (GERD) symptoms, including burning behind the breastbone, belching, and regurgitation. The burning first started 3 months ago and occurs intermittently. The burning has been quite a bit bothersome in the past week. He also has belching that occurs constantly in the past month. He also has regurgitation that occurs intermittently. He regurgitates liquids only; there is enough to fill a little of his mouth. He has not had regurgitation while sleeping in the past week. He reports hoarseness, pharyngitis, odynophagia, and shortness of breath. He does not have cough or wheezing. He tried omeprazole, and antacids and felt "somewhat better."

He does not report abdominal pain or fecal incontinence. He does not report recently having nausea, vomiting, bloating, or abdominal swelling. He is unsure if he has a neurological disorder.

He does not report blood in his bowel movements, vomiting blood, or fevers. He reports black stools for the past 3 months. He has an unintended weight loss of 20 pounds over the past 3 weeks. He has had a diminished appetite for the past 1 month. He has no history of abdominal surgeries. There is no family history of gastrointestinal cancers.

(B) Physician-generated HPI (overall impression rating in the middle third) vs. AEGIS HPI.

Physician HPI

The patient is a very pleasant 84-year-old male, who I actually saw several years ago in anticipation of a colonoscopy that he needed at that time. However, he has a new symptom complex. He now states that for the past month, he has been having postprandial symptoms of cramping pain and a sense of fullness. These symptoms occur immediately after he eats and lasts for ~10 min and then slowly resolve. It occurs usually about once a day, but it can occur with breakfast, lunch, or dinner. In between these symptoms, he has no complaints. He feels quite well and is quite active throughout the day. His appetite has been good with no significant alteration in weight. His bowel pattern is also quite regular. He denies any previous history of similar problems. In conjunction with the cramping pain that occurs after meals, he has a sense of fullness and often hears a gurgling sensation in his mid abdomen. There is no associated nausea or vomiting.

AEGIS HPI

A 84-year-old male with a history of GERD, appendectomy, and cholecystectomy who now presents with bloating. The bloating first started 4 weeks ago. He describes some abdominal distension that occurs sometimes. The distension bothers him quite a bit. He also describes his bloating as "feeling gassy." He reports passing gas about every 3–4 h. The bloating is associated with "eating food." It typically occurs 1–2 h after starting to eat. It is not associated with bowel movements.

He also reports abdominal pain. The pain first started 3 weeks ago, occurs intermittently and typically lasts for 10 min at a time. He describes the pain as cramping, says it is located in epigastrium, and reports the pain has been "a little severe" and "a little bit bothersome" in the past week. It does not radiate. It is associated with eating food. It typically occurs within 10 min of starting to eat. He is unsure how quickly the pain comes on. It is not associated with bowel movements. The pain is entirely relieved by "not eating." The pain awakens him from sleep. He reports early satiety for 4 weeks. He does not report diabetes, gallstones, pancreatitis, or peptic ulcer. He does not take aspirin or nonsteroidal anti-inflammatory drugs (NSAIDs).

He does not report diarrhea, constipation, dysphagia, or fecal incontinence. He does not report recently having nausea, vomiting, or heartburn.

He does not report blood in his bowel movements, black stools, vomiting blood, unintended weight loss, diminished appetite, or fevers. There is a family history of colorectal cancer and gastric cancer.

(C) Physician-generated HPI (overall impression rating in the lower third) vs. AEGIS HPI.

Physician HPI

A-57-year-old female with past medical history (PMH) significant for dyslipidemia, smoking, and obesity who is seen for diarrhea for 6 months and weight loss.

The patient underwent a colonoscopy 6 years ago at an outside hospital (OSH; reported normal by her, reported diverticulosis by primary care provider (PCP)).

She was treated for urinary tract infection (UTI) 6 months ago, and soon after developed 5–10 watery daily bowel movements (BMs), with 4 nocturnal episodes, associated with 40 lb weight loss and incontinence. The diarrhea improves with fasting. No specific association with dairy products or wheat. Denies abdominal pain, N/V/melena/hematemesis. Reports a single episode of bright red blood per rectum (BRBPR) after straining during defecation.

PCP work-up including stool studies was negative.

AEGIS HPI

A 57-year-old female with no Gl-related conditions who presents with diarrhea. She first began having diarrhea 10 months ago. The diarrhea has gradually worsened over time. She has loose stools 6–7 days per week. On the days she has diarrhea, she has 8–10 loose or watery stools per day. Her bowel movements are "watery, no solid pieces" (Bristol Type 7). She has urgency multiple times per day with bowel movements. She describes stool as light colored. She is not taking any medications associated with diarrhea. She does not report nausea, vomiting, bloating or swelling. She does not report a history of diabetes, irritable bowel syndrome, thyroid disease, or gastrointestinal cancer. She has not yet tried any therapy for diarrhea.

She does not report blood in her bowel movements, black stools, vomiting blood, or fevers. She reports an unintended weight loss of "40 pounds" over the past 7 weeks. She has had a diminished appetite for the past 10 months. There is no family history of gastrointestinal cancers.