Published in final edited form as:

Am J Gastroenterol. 2015 January ; 110(1): 170–179. doi:10.1038/ajg.2014.356.

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

Christopher V. Almario, MD^{1,2,3,6}, William Chey, MD⁴, Aung Kaung, MD^{1,2}, Cynthia Whitman, MPH^{1,6}, Garth Fuller, MS^{1,6}, Mark Reid, PhD^{2,6}, Ken Nguyen, MD¹, Roger Bolus, PhD⁶, Buddy Dennis, PhD⁵, Rey Encarnacion, BS⁵, Bibiana Martinez, MPH^{1,2,6}, Jennifer Talley, MSPH^{1,2,6}, Rushaba Modi, MD^{1,2,3}, Nikhil Agarwal, MD^{1,2,3}, Aaron Lee, MD¹, Scott

Corresponding Author Information: Brennan M.R. Spiegel, MD, MSHS, Director, Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Director, Health Services Research, Cedars-Sinai Health System, Pacific Theaters Building, 116 No. Robertson Blvd, 4th Floor, Los Angeles, California 90048, brennan.spiegel@cshs.org, Twitter: @BrennanSpiegel.

CONFLICT OF INTEREST

Potential Competing Interests:

None

Guarantor of the Article:

Brennan M.R. Spiegel, MD, MSHS

Specific Author Contributions:

- Christopher V. Almario, MD: Planning and conducting the study, collecting and interpreting data, drafting the manuscript, approval of final draft submitted.
- William Chey, MD: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted.
- Aung Kaung, MD: Planning and conducting the study, collecting and 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.
- Mark Reid, PhD: Interpreting data, drafting the manuscript, approval of final draft submitted.
- Ken Nguyen, MD: Planning and conducting the study, 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: Planning and conducting the study, approval of final draft submitted.
- Bibiana Martinez, MPH: Planning and conducting the study, approval of final draft submitted.
- Jennifer Talley, 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.

Kubomoto, MD¹, Gobind Sharma, MD¹, Sally Bolus, MS⁶, Lin Chang, MD³, and Brennan M.R. Spiegel, MD, MSHS^{1,2,6}

¹Division of Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, CA

²Division of Gastroenterology, VA Greater Los Angeles Healthcare System, Los Angeles, CA

³Division of Digestive Diseases, UCLA, Los Angeles, CA

⁴Division of Gastroenterology, University of Michigan, Ann Arbor, MI

⁵UCLA Computing Technology Research Laboratory (CTRL), Los Angeles, CA

⁶Cedars-Sinai Center for Outcomes Research and Education (CS-CORE)

Abstract

Background—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 computer-generated HPI. Forty-eight blinded physicians assessed HPI quality across six domains using 5-point scales: (1) overall impression; (2) thoroughness; (3) usefulness; (4) organization; (5) succinctness; and (6) 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 versus physician HPIs (3.68 vs. 2.80; $p<.001$), 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; $p<.001$), more useful (3.82 vs. 3.04; $p<.001$), better organized (3.66 vs. 2.80; $p<.001$), more succinct (3.55 vs. 3.17; $p<.001$), and more comprehensible (3.66 vs. 2.97; $p<.001$).

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

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)

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, 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. Additionally, 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) Additionally, 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 versus 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: (1) physician-documented HPIs documented in the EHR; and (2) 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 to HPIs obtained from sub-specialists 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 Patient Centered Outcomes Research Institute (PCORI) as one of eleven 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 National Institutes of Health (NIH) 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: (1) abdominal pain; (2) bloat/gas; (3) diarrhea; (4) constipation; (5) bowel incontinence; (6) heartburn/reflux; (7) disrupted swallowing; and (8) nausea/vomiting. These symptoms are based on the NIH PROMIS framework of GI symptoms, which we previously developed using data from over 2000 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 (CAT) questionnaires to measure symptom severity. The PROMIS scores are converted to a symptom “heat map” (Figure 1) that visually compares a patient’s symptoms to the general U.S. 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., B.M.R.S.). To tailor the questions to each respondent’s unique symptom experience, patients only answer a sub-set of questions conditional upon their answers. For this study, we did not time how long it took patients to complete AEGIS. However, in our prior pilot testing of AEGIS, most patients took around 10–15 minutes 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 A 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 sub-specialty fellows and internal medicine residents conduct initial evaluations in most 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 of age or older 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 non-symptomatic indications. Patients were required to read and write English and possess basic computing skills.

All patients first received usual care in their respective clinics prior to completing AEGIS. Although computer-generated HPIs would be more useful before a visit in real practice, we could not administer AEGIS prior to 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 (PDQI),(17) and blinded evaluation by a billing compliance officer.

Physician Evaluations—We recruited a panel of forty-eight reviewers from across the U.S. and Canada that included sixteen board-certified gastroenterologists, sixteen board-certified internists, and sixteen 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 ten randomly selected HPIs from ten unique patients (five 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: (1) overall impression; (2) thoroughness; (3) usefulness; (4) organization; (5) succinctness; and (6) 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.

Centers for Medicare & Medicaid Services (CMS) Level of Complexity—The level of complexity of HPIs partly determines CMS reimbursement rates for a patient encounter. CMS suggests that HPIs document a range of symptom-based elements, including: (1) location (not applicable for diarrhea, constipation, incontinence, or nausea/vomiting in this study); (2) quality; (3) severity; (4) duration; (5) timing; (6) context; (7) modifying factors; and (8) 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 vs. 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 one 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 exists 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 of a standard deviation.(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) to collect study data.(20) 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 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 versus physician HPIs. We used repeated measures analysis of covariance (RMANCOVA) 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 RMANCOVA to compare the number of CMS-recommended HPI elements present in the computer versus 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 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. Twenty-nine patients were directly recruited in the WLAVA GI clinic and completed AEGIS prior to 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 seven days after their visit (range: 2 to 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 six GI attending physicians, nine GI sub-specialty fellows, seven internal medicine residents, and one GI physician assistant – a total of 23 unique providers. Ninety-two percent (69/75) of the physician HPIs were typed directly into the EHR while the remaining 8% (6/75) were verbally dictated. No differences were seen between the typed HPIs compared to the dictated HPIs with respect to scores for overall impression ($p=.38$), completeness ($p=.47$), relevance ($p=.79$), organization ($p=.06$), succinctness ($p=.16$), or comprehensibility ($p=.06$). When comparing the 28 physician HPIs written at UCLA versus the 47 physician HPIs composed at the WLAVA, no differences were seen in overall impression ($p=.17$), completeness ($p=.26$), relevance ($p=.46$), organization ($p=.98$), succinctness ($p=.53$), or comprehensibility ($p=.98$).

When comparing HPIs written by GI attendings, GI fellows, and residents/PAs, no differences were noted in completeness ($p=.10$), relevance ($p=.14$), organization ($p=.29$), succinctness ($p=.67$), or comprehensibility ($p=.41$). However, attending physicians generated HPIs that had significantly lower overall impression scores (mean 2.50, standard

deviation [SD] 0.77) than those generated by fellows (mean 2.90, SD 0.76) and residents/PAs (mean 3.10, SD 0.44; $p=.03$). Thus, in primary and sub-group 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 to physician HPIs. Compared to the physician HPIs, blinded raters judged the computer HPIs to have significantly higher scores for overall impression (RMANCOVA adjusted $p<.001$), completeness (adjusted $p<.001$), relevance (adjusted $p<.001$), organization (adjusted $p<.001$), succinctness (adjusted $p<.001$), and comprehensibility (adjusted $p<.001$). 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 ($p<.001$). The auditor found documentation of more CMS-recommended symptom criteria in computer vs. physician HPIs with respect to the chief complaint (adjusted $p<.001$) (Table 2). More computer than physician HPIs met CMS criteria for extended status (4 elements) (100% (75/75) vs. 88% (49/56); $p=.002$). This difference lost significance when focusing only on initial visits (100% (39/39) vs. 94% (30/32); $p=.11$).

Sub-Group Analyses

In sub-group analysis focused only on initial visits ($n=39$), AEGIS HPIs were judged to be of higher overall quality (adjusted $p<.001$), more complete (adjusted $p<.001$), more relevant (adjusted $p<.001$), better organized (adjusted $p<.001$), more succinct (adjusted $p=.002$), and more comprehensible (adjusted $p<.001$). Computer HPIs maintained higher scores for overall impression (adjusted $p<.001$), completeness (adjusted $p<.001$), relevance (adjusted $p<.001$), organization (adjusted $p<.001$), and comprehensibility (adjusted $p<.001$) when excluding reviews from GI physicians (who might rate GI HPIs differently from non-GI specialists). There was a marginal effect with succinctness (adjusted $p=.08$).

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 seven 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). Ninety percent (44/49) of physician HPIs 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 to 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, Figure 3).

DISCUSSION

We studied whether a computer could gather and present patient-entered data as an HPI narrative, and compared the resulting HPIs to 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 to 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, missing important non-verbal 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 – that 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. While we cannot know from this study why there is a mismatch, computer HPIs generated pre-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 dialogue, 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 ten HPIs to each reviewer, which minimized 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. While the verbally dictated HPIs had higher scores from blinded physician raters compared to 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 to 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 provide 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, non-respondents might have low health or computer literacy. However, it is unclear whether non-responders 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 impacting 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

STUDY SUPPORT

Financial Support:

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.

REFERENCES

1. Blumenthal D. Launching HITECH. *N Engl J Med*. 2010; 362:382–385. [PubMed: 20042745]
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; 32:1478–1485. [PubMed: 23840052]
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. [PubMed: 19174390]
5. van der Lei J, Moorman PW, Musen MA. Electronic patient records in medical practice: a multidisciplinary endeavor. *Methods Inf Med*. 1999; 38:287–288. [PubMed: 10805014]
6. Sapira, JD. *The Art and Science of Bedside Diagnosis*. Baltimore: Urban & Schwarzenberg; 1990.
7. The MITRE Corporation. *Electronic Health Records Overview*. 2006
8. Asan O, P DS, Montague E. More screen time, less face time - implications for EHR design. *J Eval Clin Pract*. 2014
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; 83:225–234. [PubMed: 24380671]
10. Ventres W, Kooienga S, Vuckovic N, et al. Physicians, patients, and the electronic health record: an ethnographic analysis. *Ann Fam Med*. 2006; 4:124–131. [PubMed: 16569715]
11. Redelmeier DA, Schull MJ, Hux JE, et al. Problems for clinical judgement: 1. Eliciting an insightful history of present illness. *Cmaj*. 2001; 164:647–651. [PubMed: 11258213]
12. Spiegel BM. Patient-reported outcomes in gastroenterology: clinical and research applications. *J Neurogastroenterol Motil*. 2013; 19:137–148. [PubMed: 23667745]
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 (in press).
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
16. Bickley, L.; Szilagy, 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; 3:164–174. [PubMed: 22577483]

18. Medicare Learning Network. 1995 Documentation Guidelines for Evaluation and Management Services. 1995
19. Cohen, J. Statistical Power Analysis for the Behavioral Science. London: Academic Press; 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; 42:377–381. [PubMed: 18929686]
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; 61:344–349. [PubMed: 18313558]
22. Lucas GM, Gratch J, King A, et al. It's only a computer: Virtual humans increase willingness to disclose. *Comput Human Behav.* 2014; 37:94–100.
23. Center for Medicare & Medicaid Services. CPT/HCPCS Codes included in Range 99201–99205. 2014

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?

- 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.

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 to the general United States population with benchmarks to add interpretability to the scores, similar to a lab test.

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 prior to 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).

Figure 3.

Scatterplot of number of active GI symptoms addressed in the physician HPI versus number of active GI symptoms reported through 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.

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: *, †	
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)

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

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

* Self-reported by patient as most troublesome symptom through AEGIS.

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

TABLE 2

HPI ratings.

	Physician HPIs (N = 75)	Computer-generated HPIs (N = 75)	Adjusted p-value *
Overall impression †	2.80 (0.75)	3.68 (0.61)	< .001
Completeness †	2.73 (0.75)	3.70 (0.59)	< .001
Relevance †	3.04 (0.68)	3.82 (0.54)	< .001
Organization †	2.80 (0.80)	3.66 (0.63)	< .001
Succinctness †	3.17 (0.60)	3.55 (0.69)	< .001
Comprehensibility †	2.97 (0.79)	3.66 (0.66)	< .001
Number of CMS-recommended elements present in HPI ‡, §	5.27 (1.52)	6.05 (0.98)	< .001

Data are presented as mean (standard deviation).

CMS, Centers for Medicare & Medicaid Services; HPI, history of present illness.

* 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.

† Assessed by blinded physician reviewers.

‡ Assessed by a blinded physician compliance auditor.

§ Nineteen patients were excluded given their physician HPI had an unclear primary GI symptom.

TABLE 3

Physician HPI's primary GI symptom matched patient's self-reported primary GI symptom.*

Self-reported primary GI symptom [†]	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)

Data are presented as n (%).

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

* Limited to patients who completed AEGIS within seven days of their GI clinic visit and who reported a most troublesome symptom through AEGIS (n = 49).

[†] Self-reported by patient as most troublesome GI symptom through AEGIS.

TABLE 4

Comparing number of active GI symptoms addressed in the physician HPI versus number of active GI symptoms reported by patients through AEGIS.*

Num. of active GI symptoms reported through AEGIS	Num. of active GI symptoms addressed in physician HPI [†]
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)	-

Data are presented as mean (standard deviation).

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

* Limited to patients who completed AEGIS within seven days of their GI clinic visit (n = 55).

[†] Determined by a single blinded physician.