

Can Patient Record Summarization Support Quality Metric Abstraction?

Rimma Pivovarov, PhD¹, Yael Judith Coppleson, MPH¹, Sharon Lipsky Gorman, MS²,
David K. Vawdrey, PhD^{1,2}, Noémie Elhadad, PhD²

¹Value Institute, NewYork-Presbyterian Hospital, New York, NY; ²Department of
Biomedical Informatics, Columbia University, New York, NY

Abstract

We present a pre/post intervention study, where HARVEST, a general-purpose patient record summarization tool, was introduced to ten data abstraction specialists. The specialists are responsible for reviewing hundreds of patient charts each month and reporting disease-specific quality metrics to a variety of online registries and databases. We qualitatively and quantitatively investigated whether HARVEST improved the process of quality metric abstraction. Study instruments included pre/post questionnaires and log analyses of the specialists' actions in the electronic health record (EHR). The specialists reported favorable impressions of HARVEST and suggested that it was most useful when abstracting metrics from patients with long hospitalizations and for metrics that were not consistently captured in a structured manner in the EHR. A statistically significant reduction in time spent per chart before and after use of HARVEST was observed for 50% of the specialists and 90% of the specialists continue to use HARVEST after the study period.

Introduction

In 1999, the Institute of Medicine released *To Err is Human*,¹ which elucidated shortcomings in healthcare and ignited the movement to improve quality and patient safety. Many initiatives aimed at improving healthcare quality were developed, several of which involved the creation of large-scale data registries that are populated using both claims data and manual data abstraction².

Indicators used to assess quality of care are often buried within patient records. To accurately abstract these quality indicators, specially trained nurses manually comb through patient records to locate relevant information. Our 2,600-bed institution employs 35 full-time data abstraction specialists dedicated to reporting quality metrics for 30 databases covering 13 disease states and processes of care. Measures include CMS and Joint Commission databases for Core Measures, as well as disease-specific registries such as transplant, sepsis and stroke. The goals of the databases vary, from national benchmarking, to inclusion in value-based purchasing, to peer-comparisons for facilitating quality improvement. Participation in some databases is voluntary, such as the Society for Thoracic Surgery data registry, while participation in others is a regulatory requirement, such as the UNOS transplant database.

Each data abstracter is responsible for extracting a particular set of data using a combination of structured and unstructured clinical data including laboratory test information, comorbidities and complications through review of clinical documentation, and changes in clinical status throughout a patient's hospital stay. The complexity of each database and ease of access to information is highly variable. Some registries populate data through electronic feeds of structured documentation while others require complete manual abstraction for everything from demographics to time of symptom onset. Although every data abstracter is responsible for reporting information from patients' heterogeneous and voluminous records, the information they seek is very different, and the tools they use are different as well.

While there are many benefits in participating in quality improvement registries and databases, the burden of manual abstraction can be excessive. In fact, it is estimated that for physician practices, physicians and staff spend on average 15.1 hours per week and approximately 15.4 billion dollars annually dealing with the reporting of quality measures⁷. Despite the large time burden, it is likely that the number of reported quality metrics will continue to grow⁸.

There has been some recent success in developing natural language processing (NLP) approaches for quality metric abstraction. However, these approaches are evaluated in a retrospective fashion, outside the workflow of clinical data abstraction experts, and have largely been disease- or metric-specific. Some tools have been developed mainly for internal comparison purposes³ and others optimized to report quality metrics to specific databases: Yetisgen et al. developed a system for automated extraction of comorbidities, operation type, and operation indication for surgical care reporting⁴ and Raju et al. developed an automated method for the specific extraction of adenomas mentions⁵.

Beyond the barriers of moving from retrospective studies to interventions, and deploying NLP-based systems in a hospital setting, there exist barriers specific to the DA work: the metrics, diseases, and types of information needed vary drastically from one database to another. Furthermore, most of the DA work consists of abstracting (rather than extracting specific facts from the record) based on unstructured guidelines (although there has been work on automatically structuring quality metrics)⁶. The disease-specific work has yielded promising results and shown that some disease-specific metrics can be extracted in a fully automated fashion. Still, to date there are few available disease-specific quality metric abstractor tools as they are difficult to create and optimize, especially as the gathering of quality metrics is very heterogeneous. We hypothesize that a more holistic and broad approach could provide much needed support to a variety of DAs at once.

There has been some research demonstrating the potential benefit of holistic clinical summarization tools for physicians at the point of care^{9,10}, however, there has been much less work examining the information needs and utility of a holistic summarization tool for the purpose of quality abstraction. Our institution already provides EHR users access to HARVEST¹¹, a real-time, patient-level, summarization and visualization system. HARVEST was originally developed to aggregate relevant patient information for busy clinicians who often do not have the time to read or even skim through all available previous clinical notes. In this study, we explore the value of HARVEST in a different scenario—as a support tool for DAs in their abstraction work. HARVEST’s interface consists of three sections: a Timeline, a Problem Cloud, and a Note Panel (Figure 1). The Timeline is interactive and provides a high-level display of a patient’s inpatient, outpatient, and ED visits. The Problem Cloud displays all of the patient’s documented problems (as extracted from parsing of all the patient’s notes) with the most salient problems for this patient during the selected time range appearing larger and on top. HARVEST is a general-purpose summarization system: it extracts all problems relevant to a patient and operates on all patients in the institution. As such, it differs from the dedicated NLP approaches to quality abstraction described above in that it does not aim to do the work of the DAs, but rather support them in their search for information in the vast amount of documentation available for a given patient.

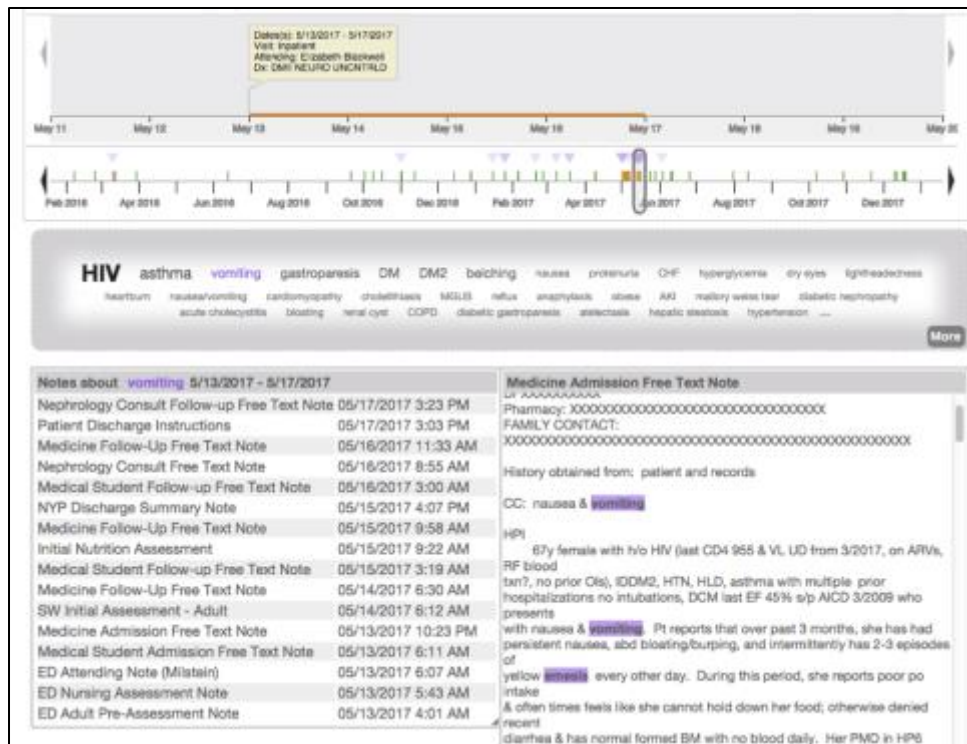


Figure 1. Screenshot of HARVEST for a de-identified patient record. The Timeline focuses on a single admission, and the user has selected "vomiting" in the Problem Cloud. The Note Panel shows the notes for that visit which mention the problem. The user has selected specific note (on the right) and all the mentions of that problem are highlighted.

Our overall research question for this study is whether a patient record summarization tool, such as HARVEST, supports the needs of DAs during their abstraction work. Towards that goal we designed a pre/post intervention study aimed at DAs with a heterogeneous range of metrics and abstraction workflows. We investigated the DAs' subjective satisfaction of the tool as well as the impact on workflow based on their usage logs of the EHR.

Methods

Data Collection

After obtaining appropriate institutional review approval, we presented HARVEST in November 2015 to Data Abstraction specialists at NYP as part of one of their "Lunch and Learn" meetings. The presentation included a 10-minute demonstration of HARVEST and where to access it on the EHR. DAs were recruited to participate in our study and consented one by one.

DA study participation consisted of four parts: (1) subjects were asked to answer a pre-study questionnaire; (2) for 4 weeks, the subjects were asked to use HARVEST during their normal course of abstraction, whenever they felt it might be useful to their abstraction and for any given patient; and (3) after the study period, the subjects were asked to respond to a post-study questionnaire. Finally, for all subjects, (4) their EHR usage logs were collected for three months before the beginning of the study and two months after the end of the study period (i.e., three months after beginning of study).

The pre-study questionnaire consisted of 20 questions and captured basic information about the DAs, including for how long they had done abstraction work, which database they abstract for, how many data elements they need to abstract for a given patient chart, and a description of their typical workflow when abstracting a chart (i.e., which part of the EHR they visit and in which order). The pre-study questionnaire also asked for expectations of where HARVEST might be most useful, and where it may fit into the DA's specific workflow.

During the 4-week study period, we sent weekly reminders to the subjects that they could use HARVEST whenever they felt it would be useful to them.

The post-study questionnaire consisted of 33 questions. The questions were inspired by the Technology Acceptance Model (TAM)¹² and the System Usability Scale (SUS)¹³. Many of the SUS questions were included in the questionnaire to measure ease of use and intention to use. Other questions were included to capture the DA's overall perception of HARVEST, for what purposes they used HARVEST, its perceived usefulness for their abstraction work, whether they had identified any unintended consequences of using HARVEST, and whether the subjects planned to continue using HARVEST in their daily work.

The questionnaires were distributed and results were collected using Qualtrics. The EHR usage logs were obtained for all systems in the NYP EHR ecosystem (which include HARVEST usage logs). The logs contained information about which function within the EHR was accessed and at what time. Examples of action types include document review, laboratory test results overview, and visit summary review. Usage log analysis enabled the computation of metrics such as time spent abstracting a given patient chart and how many patients were reviewed with or without access to HARVEST. Basic workflow information can also be derived from the usage logs as frequent sequences of actions of the different users.

Metrics

We assessed the impact of HARVEST on DA workflow in three different ways: user perception, workflow changes, and user retention. All data processing and statistical analyses were done using a combination of python and R.

User perception about HARVEST was measured through the pre and post-questionnaire questions. We compare how the DAs predicted they would use HARVEST and how they actually used HARVEST during the study period.

For the workflow analysis, we present DA self-reported workflows and measure workflow changes using the EHR log data. Workflow changes were measured in two different ways: how long DAs spend and how many EHR action types DAs take on individual patient abstractions before and after the introduction of HARVEST. The pre-HARVEST period was defined as 3 months before the DA was introduced to HARVEST, and the post-HARVEST period was defined as 3 months after the DA was introduced to HARVEST. A t-test with Bonferroni correction was used to measure whether there was a change in time spent per patient or a change in number of EHR actions for each DA before and after the introduction of HARVEST. Of note, all patients abstracted in the post-HARVEST period (whether seen with HARVEST or not) were included in the EHR log analyses.

To measure the most common workflows within HARVEST, we ran the SPADE algorithm¹⁴ for sequence event mining (using the cSPADE package in R) for all HARVEST accesses during the study period.

Retention was measured using EHR logs to see how many DAs returned to HARVEST after the study was completed and how often they used HARVEST after the study was completed. The log-based retention was compared to how the DA predicted they would use HARVEST in the future.

Results

Subjects and DA abstraction tasks

10 DAs out of the 35 DAs enrolled in the study. Subjects were nurses, and most have worked as DAs for less than 5 years (minimum experience was 1 year, maximum was 13 years). Each completed both the pre-questionnaire and the post-questionnaire 4 weeks after. The median time to completion for the pre and post-study questionnaire was 25 and 13 minutes, respectively.

Together, the enrolled subjects abstract quality metrics for 7 different databases, corresponding to 6 diseases of interest (cardiac metrics are reported to multiple databases). Databases to abstract patient information varied in the number of data elements to abstract, as well as the number of forms to fill out per patient: some have 6-10 different forms per patient while others only 1. In addition, DAs may return to the same patient over time; for example, transplant DAs fill out one form for a patient pre-transplant and another for the same patient at 6 months post-transplant.

Table 1 summarizes the self-reported DAs experiences and workflows, as captured by the pre-study questionnaire. A majority of the DAs abstract over 75 individual data elements for each patient. The data elements are found in at least 7 different EHR data types (e.g., physician and nurse notes, laboratory tests, flowsheets, medication orders, etc.). To find each of the 75+ data elements the DAs routinely visit multiple clinical systems during their workflow, likely a consequence of the fragmented and legacy systems that house patient data in our large academic medical center. One commonality across all DAs is they reported that most of their time is spent reading and abstracting from clinical notes.

Disease	Database	# DAs who abstract for disease enrolled in study	% DAs who abstract for disease who enrolled in study	Average # of data elements abstracted per patient	# different EHR data types accessed during abstraction	# different systems accessed during abstraction	Where most of DA's time is spent
Bariatric	Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program	1	100%	100+	7	3	Notes
Cardiac	Society of Thoracic Surgeons NY State DOH	3	29%	100+	8	2-4	Notes, Operative Data
Sepsis	NY State DOH	2	57%	75-100	7	2-4	Notes, Vital Signs
Stroke	Get With the Guidelines	1	25%	75-100	7	1	Notes
Surgery	American College of Surgeons	1	50%	100+	7	2	Notes
Transplant	United Network for Organ Sharing	2	35%	25-50	10	3-4	Notes, Labs, Meds

Table 1. Participant-reported data from the pre-study questionnaires on what is abstracted, from where, and what is the DA's general workflow.

When asked further about their abstraction workflow, 9 out of 10 subjects reported that finding relevant information from clinical documentation was tedious, and only 30% of DAs agreed that patient chart review was efficient. There

was in fact a very wide range of perceived time spent per patient chart review across DAs and in particular across databases. Stroke DAs reported spending on average between 20 and 45 minutes per patient; bariatric reported 35-50 minutes per patient; surgery reported 45 minutes per patient; transplant DAs reported between 20 and 90 minutes per patient; cardiac DAs reported 20-45 minutes per patient for one database and 30-120 minutes for another; and sepsis DAs reported 120-180 minutes per patient on average.

DAs Satisfaction with HARVEST

In the post-study questionnaire, all subjects agreed that HARVEST was accurate, all elements in the HARVEST interface were necessary and not redundant. All subjects felt confident using HARVEST and did not perceive any negative impact when using it for patient chart review. In addition, most subjects (80%) found the system easy to learn, easy to use, and not unnecessarily complex.

The majority of subjects thought that HARVEST positively impacted their abstraction process (60%), while the rest saw no impact on their abstraction process (40%). None saw a negative impact. Similarly, 60% of the subjects thought they would continue using HARVEST. When asked when HARVEST was useful and when it was not, subjects provided feedback as summarized in Table 2. Subjects found HARVEST most useful when patients had longer charts and they liked the problem-oriented approach to navigating the record. Subjects who did not find HARVEST helpful to their abstraction process reported that the data they abstract is typically easy to find in the patient chart already.

<p>HARVEST is most useful in when patients have longer charts</p>	<p><i>"It was helpful in all the cases I abstracted, but definitely helped me save time with longer charts, as we have many elements to abstract. It helped me narrow down timing of events such as onset of sepsis or comorbidities/complications during the patient's stay"</i></p> <p><i>"Yes! It helped me focus on areas that would have been cumbersome to abstract and review in a larger chart."</i></p> <p><i>"Definitely! and made it easier to visualize so i knew where to focus"</i></p> <p><i>"It helped me get through longer charts much more quickly."</i></p>
<p>No added value to HARVEST when the information needed is easy to access in the chart.</p>	<p><i>"Most of the patients I'm abstracting have very short visits so it's easy to read the chart through. Most of the data points are related to the current visit so it's easy to get all the information directly from the short charts. The only data points that Possibly helpful was the history. If this wasn't included in the H&P (which is rare) then I would try to use HARVEST for it."</i></p> <p><i>"I think HARVEST would be most useful for utilization review and/or data abstraction for newer registries where documentation may not be as complete and structured."</i></p> <p><i>"Although I think HARVEST is clever technology and has great potential to be useful, I will not continue to use it because there is already very specific templates and structured fields in the charts for my data abstraction."</i></p>
<p>HARVEST's problem-oriented navigation of the chart is helpful.</p>	<p><i>"Its easier to find the notes with correlating information"</i></p> <p><i>"without HARVEST, I would've had to open and skim every clinician note to find those diagnoses"</i></p> <p><i>"I use HARVEST when there is something in particular I have trouble finding."</i></p> <p><i>"I was able to use it alongside my abstraction tool to help me answer questions more quickly. There are several elements from our abstraction that HARVEST helps us speed up the search process within the chart."</i></p> <p><i>"questionable diagnoses of hepatitis and pneumonia that were confirmed by HARVEST because it was able to gather together all the source documents that contained those diagnoses"</i></p>

Table 2. Common perceptions of HARVEST from subjects and quotes from the post-study questionnaire.

Some DAs found that HARVEST played a positive role in data verification (2 found it to be very helpful, and 2 sometimes used it for verification). 6 DAs found that HARVEST was able to help identify things that they would have missed while others thought they would have found the information, but perhaps it would have taken longer.

Workflow

While the DAs were able to explain their workflow (i.e., series of actions in the EHR to carry out the abstraction for a given patient) in a detailed fashion in the pre-study questionnaire, the workflows were not as easily discernable when analyzing the EHR usage logs. Nevertheless, there were some clear patterns identified when comparing EHR usage logs 3-months pre-study, and 3-months after introduction to HARVEST.

Time spent on patient abstraction. When looking at the distribution of time spent per patient before and after HARVEST, and binning the distributions into quartiles, we found statistically significant reduction in time spent on patient abstraction for 50% of the subjects: one subject had a reduction in overall time spent over all quartiles resulting in an average 20 minutes time gained per patient, two subjects saw a reduction in time spent on short abstractions (first quartile of the distribution), and two subjects saw a reduction in time spent on long abstractions (last quartile of the distribution). For illustrative purposes, Figure 2 displays examples of distributions time before and after the introduction of HARVEST. Additionally, 1 DA saw a statistically significant increase in time spent post-HARVEST.

These findings from the EHR usage logs were confirmed by the subjects' perceived impact of HARVEST on their workflow: 4 of the 5 DAs who had statistically significant shorter times reported also reported that their workflow was shortened by HARVEST. The DA who had an increase in time spent on abstraction post study did not report a perceived increase.

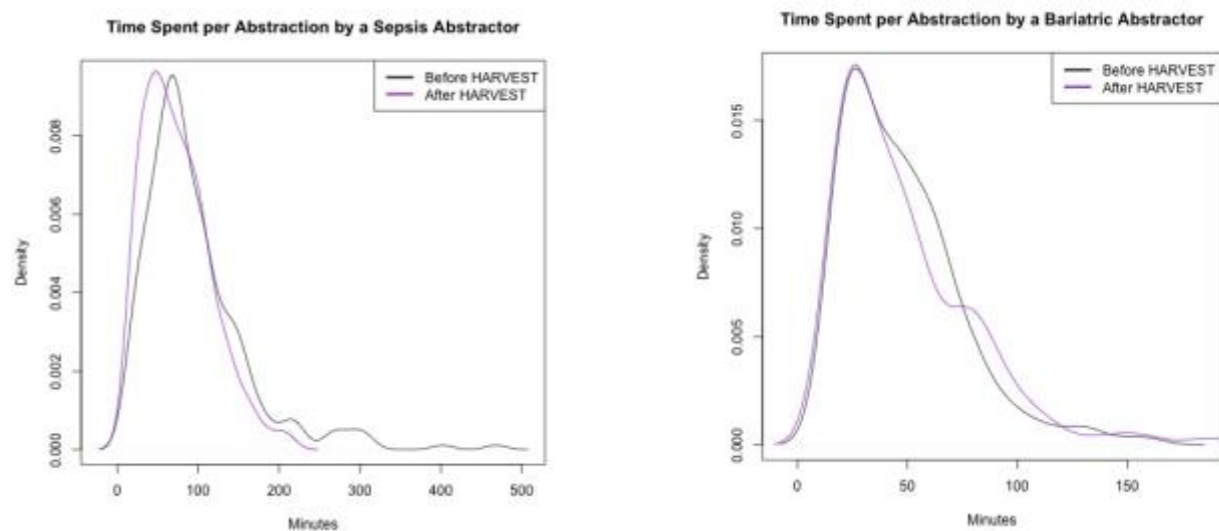


Figure 2. Density plots of the time spent per patient pre- and post-HARVEST introduction. A t-test demonstrated that the sepsis abstractor had a significant reduction ($p < 0.001$) in total time spent. On average, the sepsis abstractor saved 20 minutes per patient. The bariatric DA had no statistically significant reduction in time spent on chart review.

EHR actions. There was a significant decrease in EHR actions related to accessing the list of visits and information related to the visits. In fact, the only EHR action that increased pre- and post-HARVEST introduction was access to the HARVEST tab in the EHR. Table 3 displays some of the statistically significant differences in EHR actions and their usage pre- and post-HARVEST introduction. We hypothesize that the decrease indicates that subjects used the Timeline view of HARVEST, which visualizes all of the patients' visits from all settings, as an alternative to the traditional EHR visit list. EHR access to notes was also decreased for 3 DAs. This finding was also confirmed by the

perceived impact of HARVEST as reported in the post-study questionnaire, where access to documentation through the problem-oriented view was seen as helpful to the abstraction process.

	Average Actions per Patient	
	Pre HARVEST	Post HARVEST
Review of Documentation	95.72	63.14
	58.23	45.43
	14.57	11.75
Check Orders	17.17	7.03
	1.63	1.19
View Patient	7.76	5.6
	4.48	3.63
Access Patient Demographics	3.13	2.56
	3.64	2.84
Visit Summary	2.36	1.21
	2.68	1.13

Table 3. All EHR actions with a statistically significant pre- and post-HARVEST difference for at least two DAs. Each of the rows displayed has a Bonferonni corrected p-value < 0.001. Documentation review had a significant decrease for 3 DAs.

HARVEST workflow for patient chart abstraction. A deeper dive into the workflows within HARVEST found that aside from the basic action of loading HARVEST and setting a time period to view, the most frequent parts of HARVEST used are the selection of a problem in the Problem Cloud and opening of a note. Across the entire study period and 10 DAs, there were 222 sequences where HARVEST was used. Table 4 shows the 11 possible actions as part of the HARVEST interface, and the usage frequency of each by the subjects during the 3-month post-HARVEST introduction period.

	% of HARVEST workflows that incorporate this action
Access HARVEST tab in EHR for a patient	100%
Set Timeline to a specific date range	100%
Select problem in Problem Cloud	73%
Open Note	72%
Expand Problem Cloud to include all problems	68%
Zoom in on Timeline	62%
Select a particular visit/admission in Timeline	62%
De-select a problem in Problem Cloud	15%
De-select a particular visit/admission in Timeline	5%
Shrink Problem Cloud to most salient problems	4%
Zoom out to full longitudinal view of patient chart	0%

Table 4: The usage frequency of each of the 11 actions across all DAs in the study.

As complement to these findings from the log analysis, the pre- and post-study questionnaire showed the following trends: 80% of the subjects expected the Timeline to be useful pre-study, but only 60% found it useful after the 1-month study period. 70% of the subjects reported finding the note access functionality in HARVEST useful, which is

also reflected in the 72% usage of the “Open Note” functionality based on the EHR logs. And similarly, 60% of the subjects found the Problem Cloud useful for their abstraction workflow in the post-study questionnaire; also reflected by the 68% usage of the “Expand Problem Cloud” function in HARVEST (Table 4).

Figure 3 shows the most common sequences of HARVEST actions, as derived from all subjects in the 3-months post-HARVEST introduction. Most commonly DAs just went straight to note without clicking term. Many times they would select and re-select the timeline, move the timeline (Figure 3). Multiple “Load HARVEST for Patient” within one sequence occur if the DA accesses another tab and returns to HARVEST or refreshes the HARVEST page.

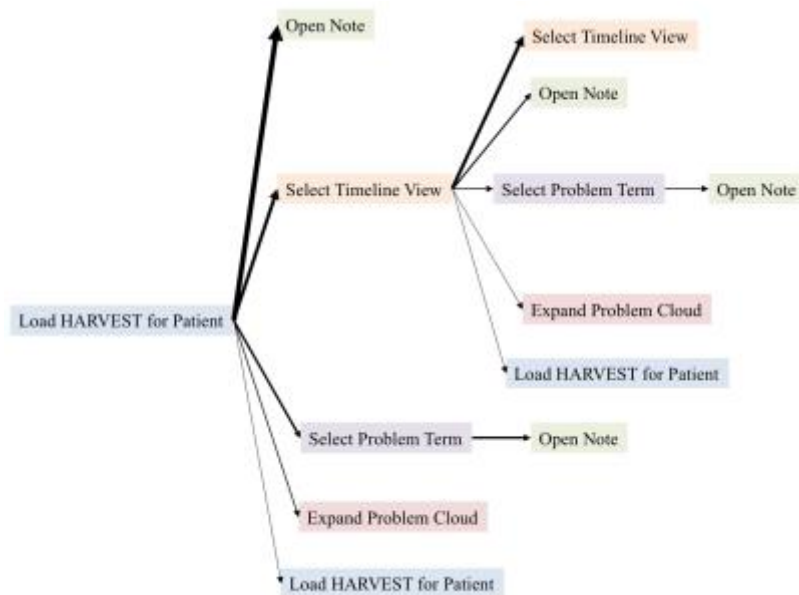


Figure 3: Most frequent workflows in the HARVEST application during one session identified using the SPADE algorithm. The width of the arrow represents frequency with which this event sequence occurs.

HARVEST Usage and Post-Study Retention

Overall, 70% of the study participants used HARVEST for at least 10% of their abstractions. The 2 Sepsis DAs used HARVEST on over 80% of their patient abstractions.

When the DAs were asked if they would continue to use HARVEST, 60% predicted that they would. Most commonly these DAs found that HARVEST would continue to be a part of their work because it helps filter and search through notes in a fast and accurate way, it has the ability to enable search for conditions and events that would be hard especially during extended admission, and it saves time. For the 40% who predicted they would not use HARVEST in the future, the problem-oriented view of the record presented by HARVEST did not align with the type of information they abstract in their tasks (for instance, one DA said “None of the problem cloud words are ones I would use.”), or that the data they needed was already easily accessible in the chart (see quotes in Table 2).

Interestingly, even though only 60% of subjects said they would continue using HARVEST, 90% of the study participants continued to use it in the post-study period of two months (see Figure 4). Even for some of the subjects for whom there was no significant time gained in patient chart abstraction, they continued using the tool for a significant portion of their charts (e.g., bariatric DA). The main exception to post-study retention were the cardiac DAs: 2 of the 3 abstractors predicted that they would continue to use HARVEST; however, according to the EHR logs their usage of the tool dropped after study completion. It is not clear exactly why the cardiac abstractors used HARVEST less than expected, but it seems that the tool did not become well integrated into their workflow. Alternatively, the sepsis team, which has one of the longest and most complex abstraction workflow (up to 180 minutes pre-HARVEST), saw such a change in workflow that they adopted the use of HARVEST as part of their protocol for chart abstraction.

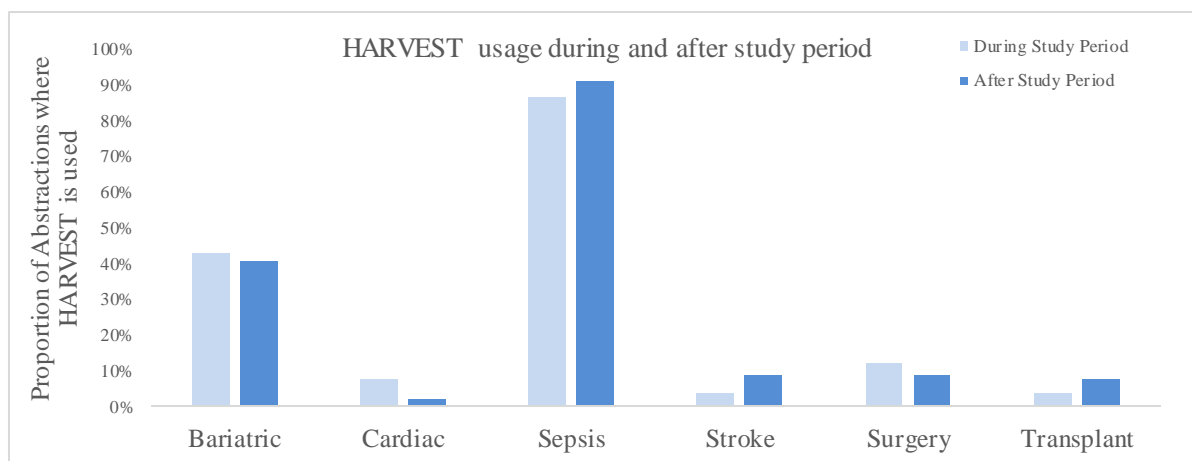


Figure 4. The retention rates of the DAs use of HARVEST. HARVEST usage stayed fairly steady after the study period for all metric types except cardiac where the use was much reduced, and stroke where the use more than doubled.

Discussion

The pre-study questionnaire reflected the complex and widely diverse activities involved in the process of chart abstraction across diseases of interest and quality databases.

Questionnaires and EHR usage log analyses provided complementary views of the impact of HARVEST on abstraction tasks across diseases. Our findings suggest that a problem-oriented patient summarizer coupled with a patient timeline can support the nurses and their information needs. HARVEST is most useful for DAs who are asked to locate data elements for patients with lengthy hospitalizations (and thus with large amounts of clinical documentation in their charts) as well as to locate data elements that are distributed in many parts of the record. However, the extent to which a summarizer such as HARVEST can be useful is variable and depends on the specific abstraction task, as well as to which extent the patient chart already documents the desired data elements in one place.

In general, DAs had a mostly positive response to HARVEST and found it useful for data verification, searching within clinical notes, and identifying the timing of events. We also found that HARVEST was able to provide statistically significant time savings for some groups of DAs. Although not all of the DAs had a statistically significant reduction in the time spent abstracting and some reported that HARVEST did not have specific terms that were necessary for their specific abstraction, 90% continued to use HARVEST after the study completion of the study.

In addition to the 10 staff who participated in the HARVEST study, the use of HARVEST continues to spread across the Division of Quality and Patient Safety at NYP. Sepsis abstractors have recently trained Trauma abstractors on how Sepsis uses HARVEST and where it fits best into their workflow. 3 DAs abstracting for trauma, cardiac, and core measures who were not enrolled in the study have been found to consistently using HARVEST. Finally, 1 DA who was recently reassigned to a new database has switched from never using HARVEST to visiting the HARVEST for 56% of their patient abstractions.

Regarding our initial question on the usefulness of a general-purpose summarizer for abstraction, our results indicate that HARVEST supports the needs of abstraction for up to 10 out of the 30 quality databases at NYP.

Limitations

Even though the DAs workflows are well-defined and are without any colleague or patient interruptions like for clinicians at the point of care¹⁵, the workflows of the DAs are still complex and unique to each. As such, the EHR usage log analysis was not successful at generalizing workflows similar to the DA self-described ones. Finally, this study is limited to one academic medical center, where HARVEST is deployed. However, it is clear that the quality abstraction burden is ubiquitous across the nation⁷.

Future Work

Following the most common suggestion from the DAs, we are working to incorporate more data types into HARVEST, presenting salient problems in clinical notes alongside laboratory test results, medications, and billing codes. This is

a non-trivial task, as we want to present all these data types in an accurate and coherent fashion, all the while keeping the problem-oriented view of the record¹⁶.

It has been suggested that HARVEST could support not only researchers conducting chart reviews, primary care clinicians at the point of care, and emergency department physicians working to obtain a brief overview of a patient they had never seen before, but also EHR users involved in other initiatives across the hospital. We hope to investigate whether HARVEST can be a useful tool when conducting root cause analyses for investigation of serious adverse events in the hospital, coding for patient billing, or chart reviews conducted by the legal department.

Conclusion

Patient chart abstraction for the sake of identifying quality metrics is a complex and time-consuming activity, yet a critical one for hospitals. Because the reporting of quality metrics requires interpretation of patient data, it is not clear that simple extraction of facts from the patient record (in particular extraction using natural language processing) is the most helpful and robust way to support data abstraction specialists. A general-purpose summarization system, such as HARVEST which supports a broad array of data abstractors in their abstraction tasks by helping them access salient part of the record, is a promising alternative solution.

Acknowledgments: This work was supported in part by a National Science Foundation award (#1344668).

References

1. Kohn LT, Corrigan JM, Donaldson MS, editors. To err is human: building a safer health system. Washington, DC: National Academy Press, Institute of Medicine; 1999.
2. Marjoua Y, Bozic KJ. Brief History of Quality Movement in US Healthcare. *Curr Rev Musculoskelet Med.* 2012;5:265-273.
3. Gawron AJ, Thompson WK, Keswani RN, et al. Anatomic and advanced adenoma detection rates as quality metrics determined via natural language processing. *Am J Gastroenterol.* 2014;109:1844-49.
4. Yetisgen M, Klassen P, Tarczy-Hornoch P. Automating Data Abstraction in a Quality Improvement Platform for Surgical and Interventional Procedures. *eGEMs* 2014; 2.
5. Raju GS, Lum PJ, Slack RS, et al. Natural language processing as an alternative to manual reporting of colonoscopy quality metrics. *Gastroint Endosc.* 2015; 82: 512-219.
6. Dentler K, Numans ME, ten Teije A, Cornet R, de Keizer NF. Formalization and computation of quality measures based on electronic medical records. *J Am Med Inform Assoc.* 2014;21(2):285-91.
7. Casalino LP, Gans D, Weber R, Cea M, et al. US Physician Practices Spend More than \$15.4 Billion Annually to Report Quality Measures. *Health Aff.* 2016;35(3):401-6.
8. Chassin MR, Loeb JM, Schmaltz SP, Wachter RM. Accountability Measures – Using Measurement to Promote Quality Improvement. *N Engl J Med.* 2010; 363(7):683-8.
9. Reichart D, Kaufman D, Bloxham B, et al. Cognitive Analysis of the Summarization of Longitudinal Patient Records. *AMIA Annu Symp Proc.* 2010:667-71.
10. Pivovarov R, Elhadad N. Automated Methods for the Summarization of Electronic Health Records. *J Am Med Inform Assoc.* 2015; 22(5):938-47.
11. Hirsch JS, Tanenbaum JS, Gorman SL, et al. HARVEST, a longitudinal patient record summarizer. *J Am Med Inform Assoc.* 2015;22(2):263-74.
12. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989;13(3):319-40.
13. Brooke, J. SUS: A “Quick and Dirty” Usability Scale. In: *Usability Evaluation in Industry*, pp. 189–194. Taylor & Francis, London (1996)
14. Zaki MH. SPADE: An Efficient Algorithm for Mining Frequent Sequences. *Machine Learning Journal.* 2001;42:31-60.
15. Holman, GT, Beasley JW, et al. The myth of standardized workflow in primary care. *J Am Med Inform Assoc.* 2016;23(1):29-37.
16. Pivovarov R, Perotte A, Grave E, et al. Learning probabilistic phenotypes from heterogeneous EHR data. *J Biomed Inform.* 2015;58:156-6.