Enhanced identification of eligibility for depression research using an electronic medical record search engine

Lisa Seyfried [a], David A. Hanauer [b,c,d], Donald Nease [e], Rashad Albeiruti [a], Janet Kavanagh [a,f], Helen C. Kales [a,f,g,*]

[a] Department of Psychiatry, University of Michigan Medical School, United States
[b] Department of Pediatrics, University of Michigan Medical School, United States
[c] Center for Computational Medicine and Bioinformatics, University of Michigan Medical School, United States
[d] Comprehensive Cancer Center, University of Michigan Medical School, United States
[e] Department of Family Medicine, University of Michigan Medical School, United States
[f] Serious Mental Illness Treatment Research Education and Clinical Center, Health Services Research and Development, United States
[g] Geriatric Research Education and Clinical Center, VA Ann Arbor Healthcare System, United States

ARTICLE INFO

ABSTRACT

Purpose: Electronic medical records (EMRs) have become part of daily practice for many physicians. Attempts have been made to apply electronic search engine technology to speed EMR review. This was a prospective, observational study to compare the speed and clinical accuracy of a medical record search engine vs. manual review of the EMR.

Methods: Three raters reviewed 49 cases in the EMR to screen for eligibility in a depression study using the electronic medical record search engine (EMERSE). One week later raters received a scrambled set of the same patients including 9 distractor cases, and used manual EMR review to determine eligibility. For both methods, accuracy was assessed for the original 49 cases by comparison with a gold standard rater.

Results: Use of EMERSE resulted in considerable time savings; chart reviews using EMERSE were significantly faster than traditional manual review ($p = 0.03$). The percent agreement of raters with the gold standard (e.g. concurrent validity) using either EMERSE or manual review was not significantly different.

Conclusions: Using a search engine optimized for finding clinical information in the free-text sections of the EMR can provide significant time savings while preserving clinical accuracy. The major power of this search engine is not from a more advanced and sophisticated search algorithm, but rather from a user interface designed explicitly to help users search the entire medical record in a way that protects health information.

Published by Elsevier Ireland Ltd

1. Introduction

Documentation and information management are fundamental aspects of patient care. Health information technology (Health IT) such as electronic medical records (EMRs), decision aids, computerized order entry and electronic prescribing have rapidly become part of daily practice for many physicians [1]. A 2006 Cochrane review concluded that Health IT

---

has been shown to significantly improve quality by increasing guideline adherence, enhancing disease surveillance, and decreasing medication errors [2]. While data on the impact of EMRs on clinical care is increasing, there is less information investigating the use of EMRs in clinical research [3].

The University of Michigan Health System (UMHS) has stored clinical information in CareWeb [4], an in-house developed electronic documentation creation and viewing system. Since 1998, UMHS has utilized CareWeb as a unified EMR. Patient encounters, problem lists and medication data are largely encoded as free text. This allows clinicians to easily and rapidly enter data, either by dictating or typing, without the constraints of a controlled medical vocabulary or predefined document structure. In 2007 alone, over 108 million lines of text representing 2.6 million clinical documents were entered into CareWeb. These patient data are not coded, making extraction of this information challenging.

Given the time and effort required for manual chart review, attempts have been made to apply search engine technology to the EMR [5–7]. Existing approaches are often not optimal given concerns with: (1) impracticality of use for patient data containing hundreds of documents (e.g. search in MS Word); (2) presentation of results in manner that makes it difficult to efficiently review medical records and keep information for each patient distinct (e.g. text processing applications such as jEdit); and (3) inability to search for words based on the case sensitivity associated with many medical terms, as well as security concerns arising from creation of an external data index (e.g. Lucene). The electronic medical record search engine (EMERSE) was created at UMHS to provide a secure and efficient way to utilize the EMR for research and clinical data abstraction [8]. EMERSE offers an intuitive user interface for searching the EMR. Search results from EMERSE are shown in a format consistent with the organization of the EMR, segregated by individual patients and separate categories for demographics, the problem summary list, clinical notes, and pathology and radiology reports. One of its most powerful features is the ability to perform batch searches across multiple patients. "Bundles", or groups of search terms, can be created to perform standardized searches of patient lists. EMERSE search bundles enable the user to search for or ignore phrases, as well as include case-sensitive searches and wildcard matches. Using the bundle, the system searches the record for these terms, and produces context-sensitive search results or "hits". EMERSE functions to protect health information; unlike manual chart reviews, records that are not of interest are ignored, not appearing as "hits".

Currently EMERSE is being used in more than 150 research studies in a variety of medical departments across our entire health system of 3 hospitals, 30 health centers and 120 clinics. Our study team is utilizing EMERSE for several research studies on later-life depression for two main purposes: (1) to quickly and efficiently screen thousands of patients for study eligibility in a manner that protects personal health information; and (2) to search or "mine" the free-text medical record to examine patient health variables that could not previously be studied in large-scale administrative data.

While we and many other study teams at our institution have found EMERSE to be invaluable, the standard in clinical research remains manual chart reviews performed by trained chart abstractors. Methods created for EMR information extraction need to be of similar accuracy and faster than manual review of medical records in order to add value.

The purpose of this study was to compare the clinical accuracy and speed of eligibility chart reviews performed using EMERSE with those done manually through the EMR. Based on our experience with this tool in several studies, we hypothesized that using EMERSE would be faster than manual chart reviews while maintaining clinical accuracy.

2. Methods

2.1. Setting

As the first step in a NIMH and IRB-approved study, we use our EMR to screen for eligibility by identifying patients who attended appointments in participating family medicine and internal medicine clinics in the prior week. Eligibility participation criteria include (1) age 60 and older, (2) primary care provider recommendation of new depression treatment within the previous month, (3) White or African American race, and (4) no history of dementia, bipolar disorder or schizophrenia. These criteria require searching numerous chart sections to confirm eligibility.

2.2. Material

For the depression study, a list of older patients who attend clinic visits are obtained weekly from an administrative database, averaging 1200 total patients from the 12 participating primary care clinics. For the purposes of the current study, a one-week convenience sample of patients (*n* = 1383) was screened using EMERSE by an experienced staff recruiter for potential eligibility. Patients' medical records contained on average 102 free-text progress notes (range 6–322).

Of these patients, 13 were determined to be eligible. An additional 37 non-eligible patients were randomly chosen from the set to form a review group of 50 cases. The study principal investigator (HCK) reviewed the eligibility decisions via manual chart review using the eligibility criteria noted above to form a gold standard for comparison. One eligible case was removed on the basis of complexity and ambiguity, leaving a sample of 49. Nine additional cases (4 eligible, 5 ineligible) were chosen from a subsequent data set for use as distractors for the second chart review. Time taken for the 9 additional records during manual review was subtracted from total time measures.

2.3. Method

The search bundle, was created from the same criteria used in manual searching. The principle investigator developed the initial bundle, which was then refined by the study team through several weekly iterations of chart screening (see Appendix A). This method for creating "bundles" has proven effective for a number of studies [9]. Fig. 1a depicts an example of a bundle search result. Raters then used search results to review individual documents (Fig. 1b) to determine eligibility.
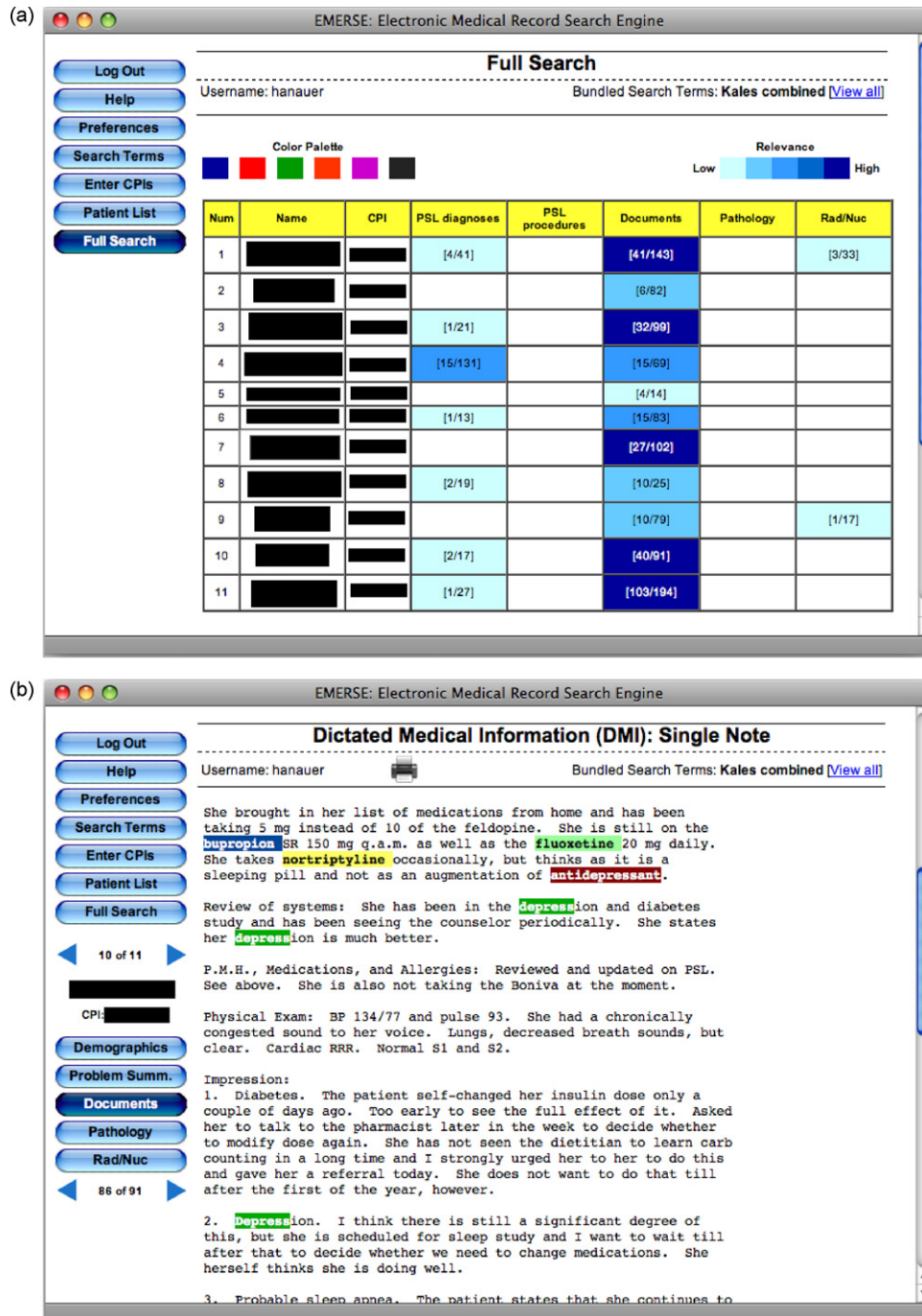
Fig. 1 – EMERSE Screenshots. (a) Screenshot of a search using the search bundle. This high-level view displays the a clinical "heat map" with rows representing individual patients and columns representing various document types in the EMR. Color-coded cells represent areas with search term "hits", the darker cells representing more "hits". Black bars were added to the figure to cover patient identifiers in order to maintain confidentiality. (b) Screenshot of an individual document. Users can click on the heat map in Fig. 1a to drill down to specific documents. Relevant search terms are highlighted and color-coded in the document to allow for rapid visual identification of search terms.

Patient records were evaluated on two occasions, by three raters with different levels of screening experience both in making study eligibility determinations and using EMERSE (8 months, 4 months, and 1 month respectively). Raters classified each case as either eligible or ineligible for the study, and recorded the time spent in the screening process. One week later, raters received a scrambled set of the same patients along with 9 distractor cases, and used a manual review of CareWeb to determine eligibility. They were instructed that the second set consisted of different patients than the first.

Clinical accuracy or concurrent validity (the measure of agreement between the results obtained by one method and the results obtained for the same population by a method acknowledged as the "gold standard") was measured by com-

Table 1 – Total time (minutes) for chart review and agreement with gold standard rater, manual vs. EMERSE.					
	Manual method			EMERSE method		
	Total time (min)	% Agreement with the gold standard	Kappa	Total time (min)	% Agreement with the gold standard	Kappa
Rater 1[*]	152	96	0.89	28	94[**]	0.83
Rater 2[*]	166	86	0.63	44	88[**]	0.65
Rater 3[*]	167	96	0.89	104	98[**]	0.94

[*] Raters are in decreasing order of experience with EMR chart review.
[**] No significant difference in % agreement of each rater with the gold standard for manual vs. EMERSE methods.

paring eligibility decisions for the original 49 cases obtained by each rater using either method (e.g. manual review or EMERSE) with the gold standard. Cohen's Kappas (measure of agreement) were used to make these pairwise comparisons. For each rater, differences in total time for chart review, manual vs. EMERSE, were compared using a nonparametric two-sample median test. All analyses were conducted using SAS Version 9.1 (Cary, NC).

3. Results

For all raters, use of EMERSE resulted in a considerable time savings regardless of level of experience (see Table 1); chart reviews using EMERSE were significantly faster than traditional manual review ($p = 0.03$). The time savings were greater for the raters with more experience (raters 1 and 2), yet even the most inexperienced rater (rater 3) saved over an hour by using EMERSE.

In terms of valid eligibility decision-making, the percent agreement of raters with the gold standard using either EMERSE or manual review was not significantly different. The majority of cases where raters showed disagreement with the gold standard resulted from falsely classifying a patient as eligible. When screening for study inclusion, these false positives will have less impact on targeted enrollment as there are typically several occasions in the process to reconfirm true eligibility.

4. Discussion

In this study, use of a medical record search engine, EMERSE, was shown to be as clinically accurate and significantly faster than manual chart abstraction for the purposes of eligibility chart reviews. Rater experience did not adversely impact level of clinical accuracy between the two methods.

These results indicate that a medical record search engine such as EMERSE has tremendous utility for reducing the time required for chart screening in clinical research both for prospective subject eligibility screening as well as for retrospective chart reviews. In addition, use of such a tool could potentially augment studies using large administrative databases with thousands of patients in order to search for variables (e.g. Mini-Mental State Exams) that are embedded in free-text chart notes and that would otherwise be impracticable to find. One of the unique features of EMERSE is that it is not fully automated, but rather augments the search capabilities of a human abstracter, which allows for

input of clinical and research experience into the process. In fact, specific document ranking techniques, such as term frequency-inverse document frequency weighting, are not used by EMERSE. As opposed to finding the "best" document with an answer to a clinical question, users must often review multiple documents, each with unique components that together help answer a clinical question. Nevertheless, the utility of using document ranking techniques should be explored in future studies.

In addition, a medical record search engine such as EMERSE also has tremendous clinical utility, particularly in urgent situations. For example, a clinician who wonders whether a patient has ever had a trial of a certain medication can do a simple electronic search, rather than manually go through multiple prior chart notes. A recent study on the use of EMRs in the Netherlands reported that over a third of practitioners chose not to search the EMR when they had questions about patients' medical issues due to time constraints [10]. An even higher percentage reported that they gave up searching for information in the EMR because it required too much time.

A growing body of literature has explored the use of information retrieval systems for health care purposes [11], in some cases to help busy clinicians access appropriate literature to answer clinical questions at the time of a patient encounter [12,13]. In contrast, little work has been done to apply such search strategies and tools to the free-text documents in EMRs [5–7]. This may be due to privacy regulations, or it may be because a different set of tools is needed. Further work should explore these issues.

This study has some limitations restricting the generalizability of the results. Both the CareWeb EMR as well as EMERSE were developed and are in use at a single institution, the University of Michigan. However, EMRs are common at many institutions and search engines like EMERSE could be adapted for those settings. Since this investigation, we have modified EMERSE for use with the EMR of the Veteran Affairs Healthcare System in Ann Arbor for use in several studies. In addition, because we standardized users to a single "bundle", we did not test each person's ability to determine an ideal set of search terms. This would have made it more difficult for us to differentiate the effects based on the users vs. the search engine itself. Among the more than 150 studies at our institution currently using EMERSE, investigators often choose to standardize the use of Bundles among team members to ensure a consistent protocol for searching the documents. Finally, with use of technology such as EMERSE, there is a chance that cases otherwise eligible might be missed due to unusual chart-note terminology or because a user might not be adept at using a

Summary points

"What was already known before our study"

- Electronic medical records (EMRs) have become part of daily practice for many physicians.
- Extraction of information contained in EMRs is needed for both clinical and research purposes.
- Extraction of clinically relevant concepts from free-text data in EMRs is complex and time consuming.
- Methods created for EMR information extraction need to be of similar clinical accuracy and faster than manual review of medical records in order to add value.

"What did our study add to our body of knowledge"

- Compared the speed and accuracy of an EMR search engine to manual medical record review.
- Results showed that using a search engine optimized for finding clinical information in the EMR could provide significant time savings while preserving the ability to make valid decisions regarding clinical eligibility.
- Relatively simple and user-friendly search tools can effectively assist researchers in accessing rich clinical data in EMRs in an efficient manner.

search engine. However, prior work has also shown the opposite. Namely, that cases missed by professional abstractors could be identified by EMERSE [8].

5. Conclusions

Using technologies such as an electronic medical record search engine to augment manual chart reviews for patient eligibility determination can result in significant time savings while preserving the ability to make valid decisions. The major power of EMERSE is not from a more advanced and sophisticated search algorithm, but rather from a user interface designed explicitly to help users search the entire medical record in a way that protects health information.

Acknowledgements

Appendix A. EMERSE search bundle used in this study

Search terms or phrases are listed, one on each line. Multi-word phrases are placed in quotes. The @ symbol is a wildcard that will match anything as long as it is part of the word to which it is attached. The ˆsymbol in front of a word means that it should be case sensitive. Custom colors can be assigned to terms to help distinguish inclusion from exclusion terms. Phrases with a - symbol in front are ignored by EMERSE which serves as a form of negation so that, for example, 'psychosis' will not be highlighted in the context of 'no psychosis'.

Inclusion	Exclusion
Phrases to find	*Phrases to find*
"major depress@"	schiz@
"depressed mood"	bipolar
ˆMDD	mani@
ˆMDE	psychosis
antidepressant	@psychotic
depress@	"mood stabilizer"
ˆSSRI	lithium
ˆSNRI	valp@
celexa	depak@
citalopram	carbamaz@
lexapro	tegretol
fluoxetine	gabapentin
prozac	neurontin
paroxetine	risp@
paxil	olanz@
zoloft	Zyprexa
sertraline	quetiapine
bupropion	seroquel
wellbutrin	haldol
deloxetine	haloperidol
cymbalta	abilify
mirtazapine	aripiprazole
remeron	demen@
venlafaxine	aricept
effexor	donepezil
trazodone	namenda
desyrel	memantine
desipramine	galantamine
norpramin	reminyl
nortriptyline	rivastigmine
Pamelor	exelon
	tacine
Phrases to ignore	cognex
"depressed heart rate"	alz@
-zyban	lewy
	"cognitive impairment"
	Phrases to ignore
	"no psychosis"
	-"no mani@"
	-"no dementia"

REFERENCES

[1] S.R. Simon, M.L. McCarthy, R.R. Kaushal, et al., Electronic health records: which practices have them, and how are clinicians using them? J. Eval. Clin. Pract. 14 (2008) 43–47.

[2] B. Chaudhry, J. Wang, S. Wu, et al., Systematic review: impact of health information technology on quality, efficiency, and

costs of medical care, Ann. Intern. Med. 144 (2006) 742–752.

[3] Y.A. Ozcan, A.S. Kazley, Do hospitals with Electronic Medical Records (EMRs) provide higher quality care? An examination of three clinical conditions, Med. Care Res. Rev. (2008), February: 1077558707313437.

[4] W.F. Bria, M.M. Shabot, The electronic medical record, safety, and critical care, Crit. Care Clin. 21 (viii) (2005) 55–79.

[5] F.F. Malamateniou, G.G. Vassilacopoulos, J.J. Mantas, A search engine for virtual patient records, Int. J. Med. Inform. 55 (1999) 103–115.

[6] W. Gregg, J. Jirjis, N.M. Lorenzi, et al., StarTracker: an integrated, web-based clinical search engine, AMIA Annu. Symp. Proc. (2003) 855.

[7] J.M. Fisk, P. Mutalik, F.W. Levin, J. Erdos, C. Taylor, P. Nadkarni, Integrating query of relational and textual data in clinical databases: a case study, J. Am. Med. Inform. Assoc. 10 (January–February) (2003) 21–38.

[8] D.A. Hanauer, EMERSE: The Electronic Medical Record Search Engine, AMIA Annu. Symp. Proc. (2006) 941.

[9] D.A. Hanauer, M.J. Engelsbe, J.A. Cowan, D.A. Campbell, Informatics and ACS-NSQIP: automated processes could replace manual record review, J. Am. Coll. Surg. 208 (January (1)) (2009) 37–41.

[10] T. Christensen, A. Grimsmo, Instant availability of patient records, but diminished availability of patient information: a multi-method study of GP's use of electronic patient records, BMC Med. Inform. Decis. Making 8 (2008) 12.

[11] F. Wiesman, A. Hasman, H.J. van den Herik, Information retrieval: an overview of system characteristics, Int. J. Med. Inform. 47 (November) (1997) 5–26.

[12] F. Magrabi, J.I. Westbrook, E.W. Coiera, What factors are associated with the integration of evidence retrieval technology into routine general practice settings? Int. J. Med. Inform. 76 (October) (2007) 701–709.

[13] P. Fontelo, F. Liu, M. Ackerman, askMEDLINE: a free-text, natural language query tool for MEDLINE/PubMed, BMC Med. Inform. Decis. Making 5 (March (10)) (2005) 5.