A Real-Time Screening Alert Improves Patient Recruitment Efficiency

Chunhua Weng, PhD[1], Candido Batres, MD[2], Tomas Borda, MD[2], Nicole G. Weiskopf, BA[1]
Adam B. Wilcox, PhD[1], J Thomas Bigger, MD[2], Karina W. Davidson, PhD[2]
[1]Department of Biomedical Informatics, [2]Department of Medicine
Columbia University, New York, NY 10032

Abstract

The scarcity of cost-effective patient identification methods represents a significant barrier to clinical research. Research recruitment alerts have been designed to facilitate physician referrals but limited support is available to clinical researchers. We conducted a retrospective data analysis to evaluate the efficacy of a real-time patient identification alert delivered to clinical research coordinators recruiting for a clinical prospective cohort study. Data from log analysis and informal interviews with coordinators were triangulated. Over a 12-month period, 11,295 were screened electronically, 1,449 were interviewed, and 282 were enrolled. The enrollment rates for the alert and two other conventional methods were 4.65%, 2.01%, and 1.34% respectively. A taxonomy of eligibility status was proposed to precisely categorize research patients. Practical ineligibility factors were identified and their correlation with age and gender were analyzed. We conclude that the automatic prescreening alert improves screening efficiency and is an effective aid to clinical research coordinators.

Introduction

Clinical research is an important step for translating basic biomedical discoveries into knowledge that will benefit clinical practice and human health. However, delays in patient recruitment can impair the ability of researchers to conduct clinical trials, the most influential form of clinical research. These delays are widespread: 86% of clinical trials are delayed in patient recruitment for up to 6 months and 13% are delayed for more than 6 months[1,2]. Recruitment delay is also expensive: in a recent large, multi-center trial, about 87 staff hours and more than $1,000 were spent to enroll each participant[3]. Inefficient enrollment can also restrict access to studies: up to 60% of patients can miss being identified to be studied[4,5]. An important factor in inefficient patient recruitment is the significant research staff time spent searching patient charts for information to compare with eligibility criteria. The rapidly expanding deployment of Electronic Health Records (EHR) invites solutions for automatically identifying potentially eligible patients using the rich data in EHR. We will refer to this process as E-screening hereafter. E-screening can exclude ineligible patients and recommend a much smaller target patient pool for manual review so that clinical research personnel can switch from "random browsing mode" to "focused and facilitated review mode" and work on only potentially eligible patients[6].

E-screening solutions have been designed for various stakeholders, including investigators, patients, and physicians, each relying on different data sources or decision support mechanisms. E-screening solutions largely fall into the following 4 categories: (1) computerized research protocol systems, such as T-Helper[7], EligWriter[8], DS-TREL[9], and OncoDoc[10], which help investigators or research coordinators to determine the eligibility of a patient using formal computable models of eligibility criteria; (2) web-based patient-enabling systems that match patients to research studies, such as caMatch[11], ASPIRE[12], TrialX[13], and ResearchMatch[14]; (3) EHR-based recruitment alerts, such as those developed by Embi et al. that enable a physician to refer a patient to researchers when that patient's EHR data meet trial eligibility criteria[15,16]; and (4) mass screening decision support for clinical researchers using clinical data repositories[6,17-21].

While these systems have great potential, their practical usefulness and feasibility has rarely been reported. For researchers to choose a particular system for a cohort study or a trial, evidence must be available about its feasibility. The benefits and drawbacks of each system need to be elucidated to better enable researchers to choose the E-screening system most appropriate for a particular study. Moreover, as Califf and others pointed out, clinical research sites are the underused component in the clinical research enterprise[22] but are critical to community-based clinical research that can reduce health disparities. Co-author Bigger and others have reported the laborious workload and inefficient workflow among overextended clinical research coordinators[23-25]. Nevertheless, enabling technologies for clinical research coordinators are still rare, as is knowledge about their unmet information and decision support needs.

We previously evaluated the use of a clinical data repository for identifying potentially eligible patients for the large randomized clinical trial ACCORD[6]. By emphasizing high negative predictive accuracy for recruiting from a highly prevalent diabetic population, our E-Screening method reduced manual review effort for clinical research coordinators by 80% without missing any eligible patients. As an extension to that study, this paper reports an evaluation of the utility and efficacy of an automated, clinical data repository-based patient prescreening alert delivered to clinical research coordinators intended to facilitate real-time patient contact and eligibility determination. In the course of conducting one of the early log analyses of informatics interventions for clinical research screening and recruitment, we also identify the most frequent factors causing patient ineligibility and the possible ratios among patients at different eligibility status. This knowledge can inform future efforts supporting clinical research feasibility assessment.

Method

1. Study Setting

We studied the real-time prescreening alerts applied to an Institutional Review Board (IRB)-approved, National Institutes of Health-sponsored observational prospective cohort study entitled Prescription Use and Lifestyle Evaluation study (PULSE), a study of post-Acute Coronary Syndrome (ACS) inpatients at the Center for Interventional Vascular Therapies of the Columbia University Medical Center (CUMC)/NewYork Presbyterian Hospital. The former is a world-renowned center with over 27,000 cardiac patients admitted annually. The hospital census data shows that about 6,100 patients are diagnosed with ACS annually. With this large study, we plan an aggressive recruitment schedule of 1,400 participants in three years or nine participants per week. The recruitment started in January 2009 and is still ongoing. Because PULSE's exclusion criteria included medical conditions, e.g., non-acute coronary syndrome (NACS), troponin levels, and EKG changes, the IRB deemed it impractical to contact and obtain consent from potential research participants before a catheterization procedure to examine these medical exclusions. Thus, we received a HIPAA waiver to screen EHR before patient contact. After medical exclusions were examined, potential research participants were approached for consent to continue eligibility screening.

2. The Informatics Infrastructure

Our informatics infrastructure for clinical research recruitment consists of two components. One is a comprehensive clinical data repository, which includes registration information, ICD-9 diagnosis, laboratory results, medication orders, and ancillary clinical notes. It is equipped with an advanced and flexible semantic integration technology, the Medical Entities Dictionary[26]. CUMC has been using the clinical data repository to support administration and research for more than 20 years[27].

The second component is a decision support system called Vigilens that is responsible for monitoring clinical events in real time and generating alerts at New York-Presbyterian Hospital (NYP). The repository feeds data to Vigilens, which was designed to complement existing hospital clinical information systems (e.g., WebCIS and Eclipsys) and currently provides a set of disparate asynchronous alerts to NYP providers. The architecture of Vigilens offers the abilities to interface with various sources of clinical data, apply both simple and complex logic to these data, and customize alert delivery to healthcare providers. Vigilens includes three core functions: (1) event manager, which interfaces with various sources of clinical data (e.g., HL7 data feed or the clinical data repository); (2) execution engine, which applies logic or rules to input data; and (3) message router, which delivers alerts to healthcare providers. Vigilens can be triggered by various clinical events, such as admit-discharge-transfer, the storage of laboratory results, the storage of reports from ancillary departments, and the processing of pharmacy orders. Vigilens can also query both coded data and narrative notes using user-supplied ICD-9 codes, medication names, diagnoses, and keywords that are present in clinical notes. Based on the events and data, Vigilens can generate user-customized alerts and deliver them via email.

While Vigilens was originally designed and used as a clinical decision support alerting system, our study used Vigilens and the clinical data repository as a recruitment tool to detect patients with ACS in real time. To facilitate real-time patient contact in concert with the workflow of the clinical research coordinators, Vigilens was configured to send three screening result emails per day at 6am, 10am, and 2pm to the two research coordinators (TB and CB). The emails did not contain any of the 18 protected health information elements specified by HIPAA except for a link to a secure, password-protected web site, where coordinators could log in to obtain patient identifiers. We used *cron*[28], the UNIX scheduler, to execute the SQL query of clinical data for patient screening regularly. The SQL statement that interrogates the clinical research repository for elevated troponin is (command to run program with input parameters: java EM/Sched_Inpat_Lab_Alert TROPONIN 68168 1 0.4):

```
query = "select me.mrn, pc.primary_time, pc.num_value, me.location" +
      " from u.management_event me, u.PROCEDUREA pa, u.PROC_COMPA pc" +
      " where me.EVENT_CODE in (32483,32467,32472,32468) and me.EVENT_STATUS = 'A'" +
   " and me.LOCATION in ('M5CC','M5GN','M5GS','M5HN','M51C','M8HN','M8HS','M9HN','M9HS','EXXX')" +
      " and me.ORGANIZATION = 1000" +
      " and me.MRN = pa.MRN" +
      " and pa.PRIMARY_TIME > me.PRIMARY_TIME" +
      " and pa.MRN = pc.MRN" +
      " and pa.PRIMARY_TIME = pc.PRIMARY_TIME" +
      " and pc.VALUE_TYPE = 1180" +
      " and pc.NUM_VALUE >= " + highRange +
      " and  pa.update_time between '" + endDate + "' AND current timestamp " +
      " and pc.comp_code in (" +
      "   select ml.des" +
      "     from u.med_lineage ml" +
      "     where ml.anc = " + medcode + ")" +
      " order by me.mrn,pc.primary_time with UR";
```

Reusing EHR data for cohort identification is a non-trivial process; heterogeneous semantic representations pose significant challenges. To identify ACS patients, the research coordinators supplied to the Vigilens software developer information about ACS symptoms (e.g., chest pain), frequent wording used by clinicians, the signs (e.g., elevated troponin), and the common ACS-specific medications, as well as the ICD-9 codes group. The developer further translated this information into the corresponding semantic representations in our EHR to construct data query logic. For example, linguistic variations such as "chest pain" and "cp" were both used to query chest pain cases in our clinical notes, and medication names were translated into Medical Entities Dictionary codes to query our structured table for medication orders.

Of note is that since the onset of this study, the research coordinators have been using an iterative process to improve Vigilens alerts precision. They review the records and, over time, accumulate knowledge about possible ways to document ACS and chest pain so that they can regularly provide such feedback to the Vigilens developer to improve its screening accuracy through the study. Therefore, rather than using machine learning, research coordinators implement a "human-computer collaboration" model to provide timely user feedback to improve Vigilens accuracy for identifying ACS patients.

3. Research Screening Workflow

The major inclusion criteria for this study were age ≥ 18, diagnosis of ACS defined as either unstable angina or myocardial infarction (MI), and physical and mental fitness to participate in the study. **Figure 1** shows our overall screening workflow, which consists of two phases: initial screening for identifying patients who meet the age and ACS criteria, followed by a coordinator interview to determine patients' physical and mental status by inquiring about more exclusion criteria, as well as their willingness to participate in the study. Initial screening identifies potential research participants and determines their medical eligibility. Medically eligible patients complete screening consent for determination of overall eligibility and safety. Completely eligible patients complete full consent and become participants in the study. Our prior studies have shown that due to missing and inaccurate data in EHR[29], not all patients identified by initial screening are eligible[6,17]. Therefore, to improve accuracy, initial screening was further divided into two sub-phases: (1) identification of a group of patients who are potentially eligible; and (2) manually reviewing the EHR of these patients to confirm their eligibility for ACS and to identify those who are approachable.

Three electronic information sources were used during initial screening: (1) lists of potentially eligible patients, generated three times a day by the automated Vigilens alerts, which automatically and regularly query the clinical data repository searching for ACS cases, emailed to coordinators; (2) daily admission lists of patients visiting CUMC's cardiac clinical units (Floor list); and (3) daily lists of patients who have undergone a catheterization (CATH) procedure that day. The first source is "pushed" to the coordinators, while the latter two contain limited information and require the coordinators to "pull" information from EHR. All the three information sources save the patient identifiers on a secure server that requires authorized access. Every day, the coordinators reviewed the EHR of those patients identified by the above three information sources to determine the presence of ACS, identify those patients who were approachable, and document each patient's eligibility status in a recruitment log. We have

previously developed a taxonomy to illustrate patient status transitions during screening and recruitment processes[17]. In this study, we extended the taxonomy to account for the two-stage consent process: consent to be screened by coordinators and consent to participate in the study.

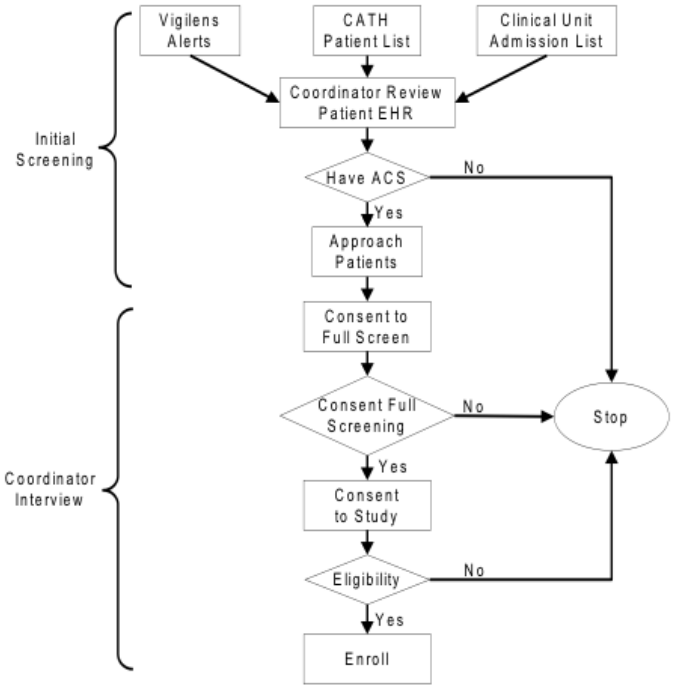

Figure 1. The patient screening and recruitment workflow.

Figure 2 shows our patient status transition diagram that includes 9 possible statuses for patients: (1) temporarily ineligible, (2) definitely ineligible, (3) potentially eligible (which also means not ineligible), (4) approachable, (5) consented to be fully screened, (6) declined to participate, (7) eligible, (8) consented to participate, and (9) enrolled.

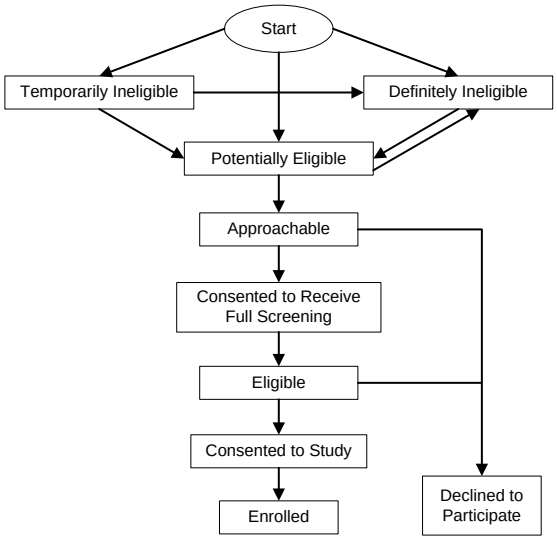

Figure 2. The patient status transition diagram.

4. Data Collection and Analysis

We obtained the recruitment log from the research coordinators (TB and CB). It comprised all patients screened between January 2009, the beginning of the study, to December 31, 2009. For each patient, the log contained basic demographic information, identification source, eligibility status, reasons for ineligibility if so, and final enrollment status. We also used the Think-aloud Protocol[30] to conduct an informal interview with the two research coordinators. Questions included: (1) "what data, knowledge, and steps are involved to confirm if a patient has ACS using EHR"; (2) "how would you rank the three information sources used during initial screening (i.e., alerts, CATH list, and clinical unit admission lists) in the order of preference"; and (3) "what is the procedure to determine eligibility after initial screening?" Triangulation of these quantitative and qualitative data allowed us to answer the following questions:

1. What is the comparative efficacy of the three methods used for initial screening: Vigilens alerts, CATH patient list, and clinical unit admission list?
2. What are the most common factors causing patient ineligibility?
3. What is the utility of real-time screening alerts for clinical research coordinator?
4. What is the duration of each step in the workflow for one patient (**Figure 1**)?
5. What additional information and decision support do research coordinators need?

We generated descriptive statistics and derived the major information needs of research coordinators. We shared these results with the research coordinators, who confirmed our interpretation.

Results

Over a 12-month period, 11,295 patients were identified by research coordinators, 1,449 were approached, 359 consented to receive full screening, 295 were enrolled, and 282 remained eligible in the study. The precision of Vigilens alerts was significantly higher than that of the patient lists search. This precision increased during the study as the research coordinators adjusted Vigilens with increasingly accurate queries. Next, we describe the overall recruitment results, the major factors causing patient ineligibility, and the coordinator feedback regarding the utility of Vigilens alerts.

1. Qualitative Study Results

Among the three screening methods, our research coordinators ranked the Vigilens alerts as the most efficient for its real-time and automatic delivery, followed by the CATH list search and clinical unit admission list search. All patients identified by the Vigilens alerts already had their information in our clinical data repository and thus were easily accessible by the coordinators to review to confirm their eligibility. In contrast, not all patients admitted to our hospital and logged in the clinical unit admission list had their complete EHR in our system, as they may have been primarily treated elsewhere; therefore, the coordinators often spent much time searching for these patients' information. Similarly, it was easier for our coordinators to contact the primary care providers of patients identified by Vigilens than by the other two information sources. Moreover, Vigilens can be easily adapted to query other cases, but not all cases had a corresponding clinical unit admission or procedure such as CATH that allowed for patient identification. The Vigilens alerts had more flexibility and extensibility than the other two sources.

2. Quantitative Study Results

Table 1 shows the 14 primary factors causing patient ineligibility across the three methods, including absence of ACS (86.87%), refusal to participate in the study (2.32%), not showing up after referral (2.07%), presence of alcohol and substance abuse (0.94%), inability to follow-up with the study (0.95%), language barriers (0.68%), cognitive problems (0.53%), psychological problems (0.36%), terminal illness (0.31%), age too old (0.04%), and incarceration (0.03%). There were no significant gender differences in each ineligibility factor except that females were less capable of following up with the study and were more likely to have cognitive or mental problems. Males were more likely to be excluded for alcohol abuse. All three methods had relatively low specificity (i.e., majority are false positives or having no ACS), although the specificity of the Vigilens alerts (19%) was more than twice that of the other two information sources (8.4% for CATH list search and 9.4% for clinical unit admission list search).

Table 1. Patients of different eligibility status across the 3 methods

Reasons to be ineligible (except for the last row)	*CATH*		*Clinical*		*Vigilens Alerts*	
1. No ACS	5,199	**91.66%**	611	**91.06%**	3,644	**80.27%**
2. Not ACS per Study	11	0.19%	0	0.00%	40	0.88%
3. Terminal Illness (e.g., cancer)	4	0.07%	3	0.45%	27	0.59%
4. Language Barriers	26	0.46%	3	0.45%	45	0.99%
5. Age (Too old to enroll)	0	0.00%	0	0.00%	4	0.09%
6. Patient Refusal	86	1.52%	10	1.49%	157	3.46%
7. Referred, Not Seen	120	2.12%	11	1.64%	94	2.07%
8. Unable to come for follow-up	36	0.63%	6	0.89%	61	1.34%
9. Alcohol or Substance Abuse	24	0.42%	3	0.45%	75	1.65%
10. Cognitive Problem	10	0.18%	2	0.30%	46	1.01%
11. Psychosis (Mental problems)	3	0.05%	4	0.60%	32	0.70%
12. Prisoner	2	0.04%	0	0.00%	1	0.02%
13. (blank)	0	0.00%	0	0.00%	45	0.99%
14. Other	37	0.65%	9	1.34%	58	1.28%
15. Eligible	114	**2.01%**	9	**1.34%**	211	**4.65%**
Subtotal	5,672	**100.00%**	671	**100.00%**	4,540	**100.00%**

Table 2 and **Figure 3** show the monthly enrollment rates using the 3 methods. Vigilens alerts generated 176 enrollments. CATH list search generated 99 enrollments. The clinical unit admission list search generated only 7 enrollments and was abandoned at the 4th month due to low specificity and efficiency.

Table 2. Monthly enrollment using the 3 Methods.

Month	*CATH*	*Clinical*	*Vigilens Alerts*	*Total*
1			1	1
2		1	3	4
3	6	2	9	17
4	18	4	17	39
5	5		12	17
6	12		13	25
7	9		26	35
8	7		24	31
9	12		16	28
10	11		21	32
11	11		22	33
12	8		12	20
Total	**99**	**7**	**176**	**282**

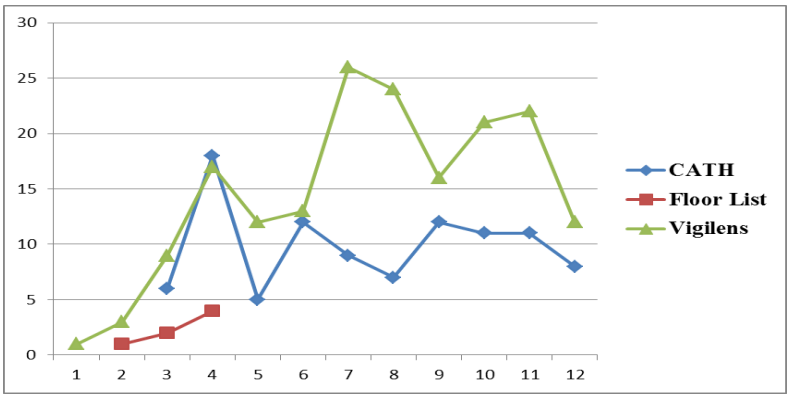

Figure 3. Monthly enrollment using the 3 methods.

Throughout the study, the Vigilens alerts were consistently more efficient than the other two information sources. The average time to process an alert was 2 minutes. This included logging into the secure web site and performing a patient-based EHR review for chest pain and ACS. In contrast, use of the CATH patient lists or clinical units required multiple steps, including identifying patients by name or medical record number, searching for patient medical records (e.g., the records of some patients who visited our clinic may not be electronically available in our system), locating the documents describing ACS symptoms for patients with this condition, and creating electronic notes for those patients. The average time to complete these steps was about 15 minutes per patient. As shown in Table 2, the research coordinators identified 114 (2.01%) eligible patients among the 5,672 in the CATH lists and identified 211 (4.65%) eligible patients among the 4,540 recommended by the Vigilens alerts.

Table 3 shows the distribution of primary factors causing patient ineligibility over time. Among 1,449 patients who had ACS, 282 (25%) were eligible for the study; the remaining 75% were ineligible for one of the primary reasons before or after initial screening. 17.9% patients explicitly refused to participate and 15.9% did not appear after the referral, which can be attributed to either implicit patient refusal or missed recruiting opportunities by the research coordinators due to time constraints or workflow factors (**Figure 1**). Overall, 34% patients were eligible but not successfully recruited to this study; further studies are needed to investigate the reasons for this. About 25% of eligible patients consented to be fully screened, among whom 82% enrolled in the study. "Patient refusal", "patient referred but unseen", and "alcohol and substance abuse" were the most prevalent criteria among patients who were eligible but not part of the study. A few patients were ineligible because they were prisoners or too old.

Discussion

In this paper, we report on the evaluation of a real-time patient identification email alert applied to a large clinical study. We triangulated qualitative (i.e., informal interviews with research coordinators) and quantitative (i.e., recruitment log analysis) data to compare the efficacy of the alert delivered to research coordinators (the "push" mode) with that of traditional case search methods in clinical units (the "pull" mode). Our study confirmed that the automated patient identification alert is much more efficient and preferred by the research coordinators than CATH lists or patient admission lists. The alert also holds great potential for further improvement.

This study elucidates patterns of patient ineligibility factors. This is the first study to characterize patient eligibility status in different categories, such as "potentially eligible", "approachable", "consentable", "eligible", and ultimately "enrolled". This taxonomy provides greater nuance than conventional binary classification and enables research coordinators to more precisely calculate the ratios of patients of different eligibility status during screening. Although the generalizability of the ratios is untested, we hope this case study contributes to the development of reference standards for supporting clinical research feasibility studies. Our eligibility status taxonomy can also help others categorize their patients with different eligibility status for measuring screening efficiency in future studies.

When analyzing Vigilens errors, we identified an interesting factor that causes patient ineligibility: not ACS per study criteria. A patient considered an ACS case by local hospital criteria might not be considered to have ACS as defined by the protocol-specific research eligibility criteria. After interviewing the research coordinator, we learned that patients in this category had "elevated troponin" in the notes; however, the hospital criteria set a threshold of ">0.09" for elevated troponin, whereas the protocol's threshold is ">0.36". Therefore, a particular semantic representation can have different meanings in different contexts, which can adversely affect screening.

There are a number of limitations to this study. First, we only studied inpatient settings. There might be different recruitment challenges in other patient care settings, such as emergency room or outpatient clinics. Further studies are needed to assess the efficacy and utility of real-time patient identification alerts in different types of clinical research studies in various patient care settings. Second, this informatics observational study was a retrospective analysis of the clinical study and so was not prospectively designed to test the relative efficiency of these screening methods. Our data collection and analysis were retrospective; therefore, we did not capture all information systematically and, instead, had to rely on human recall and qualitative methods. Third, the three data sources were not mutually exclusive. Therefore, a patient was identified sometimes by more than one method or often by the same method. When identified by more than one source, the patient was assigned a source arbitrarily in the recruitment log. The research coordinators also sometimes received more than one Vigilens alert for the same patient; the burden of redundancy was high especially when the inpatients had a long length of stay. There was no tool available to help the research coordinators reconcile information from the three data sources, which resulted in much unnecessary and tedious work for the coordinators to identify duplicates manually. Future efforts to improve the alerts should consider using a lookup table listing all screened patients using all the data sources and should avoid sending multiple alerts for the same patient. Research coordinators often work with multiple documents

simultaneously and an information reconciliation tool would be valuable. Finally, we did not know how research coordinators distributed their time among the three methods. A future prospective study with more systematic data collection may more completely answer research questions raised in this study.

Research coordinators should play a major role in increasing the accuracy of alerts. In this study, we only used keyword-based content match without support for disambiguation or negation detection; therefore, our approach did not have the natural language processing (NLP) capacity. A future improvement would be adding NLP components. The coordinators started with simple keywords, which did not fully capture all the ACS cases. As they reviewed each case and accumulated more knowledge about how ACS was documented in EHR, the coordinators became more proficient in selecting terms to help with query expansion and thus significantly improved Vigilens accuracy over time. We obtained the latest recruitment report as of February 14, 2011 and found the Vigilens query precision increased from 12.5% in January 2009 to 82% in February 2011. We intend to study what queries helped with the improvement in more detail to inform accurate EHR phenotyping research. Moreover, this improvement itself is an exciting result. Although we were unable to fully evaluate the alerts accuracy improvement in our current study, the improvement has encouraged us to attempt to build a "human-computer collaboration" model for informatics intervention development. This improvement demonstrates the importance of research coordinators in the effort to design a more efficient recruitment process. They can provide valuable feedback to improve the performance of an informatics intervention. In our future work to improve Vigilens, we will provide an informatics infrastructure to more proactively support research coordinators and incorporate their feedback in a more efficient, timely, and user-friendly fashion.

Conclusion

This study used a retrospective data analysis method to compare the efficacy of three methods for clinical research screening. Our results demonstrate the potential of an automated screening alert linked to a clinical data repository for increasing the efficiency of the clinical research recruitment process. We elucidated practical factors causing patient ineligibility and identified no gender differences in most of them. We also contributed preliminary information about the percentages of patient populations across the spectrum of eligibility statuses. This can be valuable for assessing the population size required for a recruitment goal and for building informatics models enabling clinical research feasibility studies.

Acknowledgements

This research was funded under NLM grant R01LM009886 and R01LM010815, CTSA award UL1 RR024156, NHLBI grants HL088117, 101663, and AHRQ grant R01 HS019853. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NIH.

References

1. Sullivan J. Subject Recruitment and Retention: Barriers to Success. *Applied Clinical Trials.* April 2004 2004.

2. Sinackevich N, Tassignon J-P. Speeding the Critical Path. *Applied Clinical Trials.* January 2004 2004.

3. Rubin R, Rujimoto W, Marrero D, Brenneman T, Charleston J, Edelstein S. The Diabetes Prevention Program: recruitment methods and results. *Controlled Clinical Trials.* 2002;23:157-171.

4. Carlson R, Tu S, Lane N, et al. Computer-based screening of patients with HIV/AIDS for clinical-trial eligibility. *Online J Curr Clin Trials.* 1995.

5. Fink E, Kokku PK, Nikiforou S, Hall LO, Goldgof DB, Krischer JP. Selection of patients for clinical trials: an interactive web-based system. *Artificial Intelligence in Medicine.* 2004;31(3):241-254.

6. Thadani SR, Weng C, Bigger JT, Ennever JF, Wajngurt D. Electronic Screening Improves Efficiency in Clinical Trial Recruitment. *Journal of the American Medical Informatics Association.* November 1, 2009 2009;16(6):869-873.

7. Musen MA, Carlson RW, Fagan LM, Deresinski SC. T-HELPER: Automated Support for Community-Based Clinical Research. Paper presented at: 16th Annual Symposium on Computer Applications in Medical Care, 1992; Conference; Location. 719-723.

8. Gennari J, D. Sklar, Silva J. Cross-Tool Communication: From Protocol Authoring to Eligibility Determination. Paper presented at: Proc of AMIA'01, 2001; Washington DC. 199-203.

9. Breitfeld PP, Weisburd M, Overhage JM, Sledge G, Jr., Tierney WM. Pilot Study of a Point-of-use Decision Support Tool for Cancer Clinical Trials Eligibility. *JAMIA.* November 1, 1999 1999;6(6):466-477.

10. Seroussi B, Bouaud J. Using OncoDoc as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials. . *Artificial Intelligence in Medicine.* September 2003 2003;29(1):153-167.

11. Cohen E, al. e. caMATCH: A Patient Matching Tool for Clinical Trials Paper presented at: caBIG annual meeting 2005.

12. Niland J. ASPIRE: Agreement on Standardized Protocol Inclusion Requirements for Eligibility. *Unpublished work.* 2007.

13. Trialx. http://www.trialx.com Accessed 2/27/2011.

14. Harris PA, Lane L, Biaggioni I. Clinical research subject recruitment: the Volunteer for Vanderbilt Research Program www.volunteer.mc.vanderbilt.edu. *J Am Med Inform Assoc.* Nov-Dec 2005;12(6):608-613. 1294031

15. Embi PJ, Jain A, Clark J, Bizjack S, Hornung R, Harris CM. Effect of a Clinical Trial Alert System on Physician Participation in Trial Recruitment. *Arch Intern Med.* October 24, 2005 2005;165:2272-2277.

16. Embi P, Jain A, Clark J, Harris C. Development of an electronic health record-based Clinical Trial Alert system to enhance recruitment at the point of care. Paper presented at: AMIA Annu Symp Proc 2005. 231-235.

17. Weng C, Bigger J, Busacca L, Wilcox A, Getaneh A. Comparing the Effectiveness of a Clinical Data Warehouse and a Clinical Registry for Supporting Clinical Trial Recruitment: A Case Study. Paper presented at: Proc of AMIA 2010 Fall Symp., 2010. 867-871.

18. Weber S, Lowe H, Malunjkar S, Quinn J. Implementing a Real-time Complex Event Stream Processing System to Help Identify Potential Participants in Clinical and Translational Research Studies. Paper presented at: AMIA Annu Symp Proc 2010; Conference; Location. 472-476.

19. Dugas M, Lange M, Müller-Tidow C, Kirchhof P, Prokosch H-U. Routine data from hospital information systems can support patient recruitment for clinical studies. *Clinical Trials.* April 1, 2010 2010;7(2):183-189.

20. Grundmeier RW, Swietlik M, Bell LM. Research subject enrollment by primary care pediatricians using an electronic health record. *AMIA Annu Symp Proc.* 2007:289-293.

21. Embi PJ. Increasing Physician Participation in Clinical Trial Recruitment. *Physicians Weekly.* 2006;23(3).

22. Califf RM. Clinical Research Sites—The Underappreciated Component of the Clinical Research System. *JAMA: The Journal of the American Medical Association.* November 11, 2009 2009;302(18):2025-2027.

23. Khan S, Kukafka R, Payne P, Bigger J, Johnson S. A day in the life of a clinical research coordinator: observations from community practice settings. *Stud Health Technol Inform.* 2007;129(Pt 1):247-251.

24. Khan SA, Kukafka R, Bigger JT, Johnson SB. Re-engineering Opportunities in Clinical Research using Workflow Analysis in Community Practice Settings. Paper presented at: AMIA Annu Symp Proc 2008. 363–367.

25. Khan SA, Payne PRO, Johnson SB, Bigger JT, Kukafka R. Modeling Clinical Trials Workflow in Community Practice Settings. Paper presented at: AMIA Annu Symp Proc., 2006. 419–423.

26. Cimino JJ. The Medical Entities Dictionary. http://med.dmi.columbia.edu/index.html. Accessed October 5, 2009.

27. The Clinical Data Warehouse. http://ctsa.cpmc.columbia.edu/rdb/request.html Accessed March 17, 2011.

28. CRONTAB. http://pubs.opengroup.org/onlinepubs/009604599/utilities/crontab.html. Accessed July 13, 2011.

29. Botsisa T, Hartvigsena G, Chen F, Weng C. Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. Paper presented at: Proc of Clinical Research Informatics Summit 2010. 1-5.

30. Kushniruk AW, Patel VL, Cimino JJ. Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces. *Proc AMIA Annu Fall Symp.* 1997:218-222.

Table 3. The summative monthly multi-source report of patients at different eligibility status

MM/YY	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*
1/09	10	1	9	0	0	1	0	1	3	1	0	2	0	0	1	0	0	0	0	1	1	0
2/09	711	622	89	0	2	1	0	15	13	13	0	13	0	4	28	3	2	4	1	18	3	15
3/09	1,245	1,130	115	0	4	8	0	20	19	9	6	34	11	21	20	1	1	1	1	16	2	14
4/09	1,252	1,102	150	0	2	13	0	20	33	11	12	11	1	8	39	2	1	1	3	32	3	29
5/09	1,196	1,022	174	0	5	4	0	41	33	15	15	13	0	15	33	1	4	2	2	24	2	22
6/09	1,084	941	143	0	2	9	0	26	35	16	14	9	0	5	27	0	1	1	0	25	1	24
7/09	982	857	125	5	1	8	0	21	17	9	12	8	0	7	37	4	1	0	4	28	0	28
8/09	1,068	934	134	7	1	5	1	18	22	11	7	11	0	7	44	1	3	2	2	35	0	35
9/09	928	781	147	9	6	6	1	25	14	10	15	16	0	16	29	1	2	0	1	25	0	25
10/09	939	816	123	5	4	11	0	26	13	8	5	10	0	9	32	1	1	1	0	30	0	30
11/09	1,075	938	137	13	4	4	0	25	17	6	10	7	0	7	44	3	1	0	1	39	0	39
12/09	805	702	103	11	4	5	2	22	11	6	4	5	1	7	25	0	0	1	2	22	1	21
Total	11,295	9,846	1,449	50	35	75	4	260	230	115	100	139	13	106	359	17	17	13	17	295	13	282
	1	87.2%	12.8%																			
			1	3.5%	2.4%	5.2%	0.3%	17.9%	15.9%	7.9%	6.9%	9.6%	0.9%	7.3%	24.8%							
															1	4.7%	4.7%	3.6%	4.7%	82.2%		
																				1	4.4%	95.6%

Column headers:

1. Patients screened using all data sources
2. Not Acute Coronary Syndrome (NACS)
3. Acute Coronary Syndrome (ACS)
4. Note Acute Coronary Syndrome per study
5. Terminal illness
6. Language barrier
7. Unfit age group
8. Patient refusal to participate
9. Patient referred but unseen
10. Unable to follow-up with the patient
11. Substance abuse
12. Mental illness
13. Prisoner
14. Other
15. Consent to be screened
16. Mental illness identified after consent
17. Substance abuse identified after consent
18. Unable to complete the study
19. Other
20. Enrolled
21. Post-enrollment exclusion
22. Eligible

**: The total numbers of patients in Table 3 were slightly higher than that in Table 2 because it includes some rarely used data sources. The difference did not change the percentages of patients in each status category*