

Information Technology in PBRNs: The Indiana University Medical Group Research Network (IUMG ResNet) Experience

Abel Kho, MD, MS, Atif Zafar, MD, and William Tierney, MD

Objective: Research in practice-based research networks (PBRNs) is hampered by difficulty managing, identifying, and enrolling potential subjects. Well-designed informatics applications can greatly improve these processes.

Methods: We considered a literature review, discussion with PBRN researchers, and personal experience to outline important principles to apply when considering electronic data collection in a PBRN. We provide specific working examples of electronic means we use to improve data collection and patient enrollment.

Results: Our PBRN has screened more than 18,000 patients and enrolled more than 6000 study subjects in 5 years. Less than 2% of potentially eligible patients are missed by our research assistants. We achieved this high rate of success through extensive integration of the ResNet infrastructure (research databases and personnel) with an electronic medical record (EMR) system and computerized provider order entry. We make extensive use of widely used standards for data storage, definition, and transmission to ensure data reusability. We successfully implemented a real-time means to identify follow-up patients.

Conclusion: Electronic data collection can greatly facilitate PBRN research, particularly by improving data management and identification of eligible patients. Key principles to ensure successful implementation include use of data standards and centralized electronic data management. (J Am Board Fam Med 2007;20:196–203.)

Performing studies in real-world primary care practices is hampered by difficulty in identifying and enrolling potential subjects. Despite the best intentions, the busy clinician may lack the time to identify patients that are eligible for a research protocol. This has resulted in few subjects being recruited from nonacademic practices (eg, only 3%

of patients with cancer being cared for by community oncologists are entered into clinical trials).¹

Even when physician time for recruiting and enrolling patients is not a significant barrier, the process of navigating complex algorithms on paper is error prone.² There is additional cost incurred with transcribing the acquired data from paper to electronic forms for data analysis. Given these drawbacks of paper-based data collection, electronic data collection strategies has advantages for reducing some of these problems. We describe the informatics innovations at ResNet, a practice-based research network (PBRN), as a successful working example and provide specific recommendations for PBRNs considering adopting electronic data collection to facilitate research.

Background

Health information technology (IT) consists of a complex interplay of technologies, policies, standards, and users. However, understanding just a handful of key principles and terms provides clinicians and PBRN administrators with a starting

This article was externally peer-reviewed.

Submitted 7 July 2006; revised 10 October 2006; accepted 12 October 2006.

From Regenstrief Institute, Indianapolis, IN (AK, AZ, WT); General Internal Medicine, Northwestern University, Chicago, IL (AK); and Indiana University School of Medicine, Indianapolis, IN (AZ, AK, and WT).

Conflict of interest: none declared.

Funding sources: AK was supported in part by a grant from the National Library of Medicine (T15 LM007117). AZ and WT are supported in part by Contract HS-230-02-0008 from the Agency for Healthcare Research and Quality (AHRQ).

Previous presentation: This article is based on a presentation made at the 2006 Agency for Healthcare Research and Quality National Practice-based Research Network Conference, Bethesda, MD, May 15–17, 2006.

Corresponding author: Abel Kho, MD, MS, General Internal Medicine, Northwestern University, 676 N. St. Clair Street, Suite 200, Chicago, IL 60611 (E-mail: abel.kho@nmff.org).

point for planning and discussions to start collecting data electronically.

Store Data in a Consistent Format, Using Widely Accepted Standards

Standards exist for coding much of clinical medicine. Examples include the International Classification of Diseases Ninth Edition (ICD-9) system for coding diagnoses and the Current Procedural Terminology (CPT) system for coding procedures. Logical Observations Identifiers Names and Codes (LOINC codes; www.regenstrief.org/loinc/) describe a widespread system for coding laboratory studies.³ The National Library of Medicine has purchased licensing for SNOMED codes (Systematized Nomenclature of MEDicine)⁴, which cover a broad range of clinical conditions. Each of these standards provides a common and machine-interpretable vocabulary for person-machine and machine-machine interactions. “Mapping” existing data to these widely accepted standards (creating electronic matched lists of existing data to a standardized vocabulary) will take time up front but can save time later during data collection.

Understand How to Define Disease States and Conditions

There are many ways to define disease states. How these are represented in your system will be of critical importance as you extract data and identify patients from enrollment criteria. For example, diabetes can be defined in any of the following ways: fasting blood sugar >126, random blood sugar >200, person on insulin or other diabetic drug, person with a diagnosis of diabetes (ICD 250.XX), person with a Hgb A1c >8 or other value, etc. The definition you choose to use will affect extraction of data for determining inclusion and exclusion criteria. Using multiple criteria for describing a disease state may prove useful for future research.

Transmit Data in a Standardized Format

If your institution uses an EMR or has access to electronic data streams from lab facilities or hospital systems, then you may be able to tap into these streams to build a repository dynamically using standard messaging protocols. For example, Health Level Seven (HL7) describes a standard for transmitting textual health information.⁵ Proper use of the HL7 standard guarantees that when information is transmitted, the same data elements (ideally

```
PID||PATID1234^5^M11||DOE^JOHN^A||19610615|M||C|
1200 N ELM STREET^ANYWHERE^AL^35044|
GL|(800) 5553434|
|S||PATID12345001^2^M10|123456789|987654^AL|<cr>
```

Figure 1. A sample HL7 message. Pipes (|) separate discrete data elements.

much of which is coded!) will be found in the same parts of each message. For example, a sample registration HL7 message may appear as shown in Figure 1. Although not readily human interpretable, the important principle is that for HL7 compliant messages, we can always count on certain data elements to consistently show up in the same location in an electronic message. For example, in Figure 1, the patient ID number appears in the fourth segment, and the date of birth appears in the eighth segment. This consistency can be a great boon for the researcher interested in collecting electronic data.

A nearly universal type of HL7 message is the ADT message (admission, discharge, transfer). ADT messages encode information such as patient demographics, some clinical diagnoses, and registration data. ADT messages can be tapped to build up a patient registry using an interface engine. An interface engine is a software tool that decodes HL7 and other clinical message types, parses them, and stores the data into a repository automatically. In our example, we adopted a simplified open source interface engine to receive incoming electronic messages.

Use Tools That Are Familiar in Your Environment

It is far simpler (and less costly) to use what already exists within a particular environment rather than embedding new technology into busy workflows. Whatever the types of computers available (eg, desktops, laptops, tablet PCs, or PDAs), it makes sense to use these for collecting research data (prompts/alerts, simple questionnaires, etc) from the clinicians who are already familiar with their operation at the point of care. In our example, we used the existing desktops used by clinicians for routine care.

Pool the Skills of Team Members

A typical team may consist of a mix of programmers and IT experts, data managers, a study coordinator, and research assistants (RAs). The role of the IT

expert is to maintain and administer data servers, program and maintain the repository, keep your network running, and maintain data security and HIPAA compliance. The role of the data manager is to extract data from the database. Often the role of the IT expert and data manager can be taken on by the same person, given sufficient expertise. The role of the study coordinator is to oversee the RAs and help troubleshoot and operate the trials. Although personnel costs will likely form the greatest proportion of PBRN infrastructure costs,⁶ the success of an IT rollout is critically dependent on the strength of the people driving the implementation. Centralizing tasks electronically has improved our efficiency and allowed a small research team to manage a wide range of clinics.

ResNet: A Working Example

So what might a successful implementation look like? Using the example of ResNet, we describe how a focus on electronic data management and electronic data collection strategies has enabled successful and efficient resource usage in our PBRN. We describe the evolution of our research data management from paper to computer and demonstrate the role that each of these principles, from standardized data storage and messaging to our centralized team approach, play in this ongoing work in progress.

Established in 1999, ResNet consists of 17 primary care sites serving a diverse population of insured and uninsured patients throughout Indianapolis. More than 110 physicians (excluding residents) practice at these sites, about half of whom practice there every day. These physicians treat approximately 100,000 patients per year (two thirds are children) in more than 300,000 visits annually.

ResNet effectively uses informatics tools to streamline patient recruitment and follow-up. It is built on the Regenstrief Medical Record System (RMRS), a comprehensive EMR system, developed over 3 decades of routine use for clinical care.⁷ The RMRS makes extensive use of standards, including LOINC for laboratory tests and CPT and ICD-9 codes for diagnoses. To date, the RMRS contains more than 1.5 million patients and more than 800 million separate and standardized observations. These data are retrieved and managed by the Regenstrief Data Core, which consists of 8 full-time master's-level data managers housed at the Regenstrief Institute.

In the 5 years since its inception, ResNet has recruited subjects for more than 30 studies with more than \$15 million in extramural direct cost funding. For these studies, ResNet RAs have screened more than 18,000 primary care patients, enrolling more than 6000 subjects (66% of those patients found to be eligible by screening criteria). Importantly, less than 2% of potentially eligible patients keeping scheduled appointments were missed by the ResNet RAs. ResNet has achieved this success through the following informatics innovations.

Centralized Electronic Data Management

Before a study is operationalized, data managers extract from the RMRS a list of eligible patients based on the inclusion criteria as provided by the principal investigators. Once an eligible list of patients has been extracted, it is provided to the clinicians who authorize the study for their patients. Letters are sent to the patients informing them of the study. Subsequently, the registration records (IDX) are queried to determine when the patients are due for an appointment. Daily lists of patients are then generated for the various ongoing studies.

Before ResNet, patient lists were printed on paper, and the same paper lists were used to inform the data managers which patients were contacted, ineligible, refused, enrolled, etc. Multiple files (1 per study) were maintained on the ResNet computer system containing data on approached, enrolled, and ineligible patients. Subsequent studies avoided contacting patients enrolled in ongoing projects by extracting data from each of these separate files.

This entire process is now electronic and consolidated into a single, centrally maintained list. The weekly lists now contain *all* of the patients eligible for *any* ongoing study with scheduled appointments along with all the studies for which they are eligible. These studies are prioritized based on which study began first, although occasionally other criteria govern the order of studies into which the patient is recruited.

This prioritized list is sent to the ResNet RAs' computers electronically as a Microsoft Access database. The RAs enter recruitment information about all patients contacted into this database, and then the database is sent back to the data manager. The data manager then merges the data into a master database that is used to generate future lists.

ResNet Physician Checklist

PI: Dr. Bill Tierney
Title: Ethnic Differences in Health Problem Reporting in Depressed Patients

The following patients are potentially eligible for a study looking at ethnic differences in health beliefs of patients with depression. This is a survey study in which patients will be asked to answer questions regarding demographic information, health status, health beliefs and health knowledge at baseline, 12 months, and 24 months. Please indicate below which patients may be approached by a ResNet Research Assistant for potential participation in this study. Please return this list via campus mail in the attached self-addressed envelope or give to your ResNet Research Assistant. If you have any questions regarding this study, please contact [name], the ResNet Program Manager, at [number]. Thank you!

CLINICIAN: Abel Kho
CLINIC: PCC

AFTER REVIEWING THIS LIST, RESNET HAS PERMISSION TO APPROACH FOR THIS STUDY:
(Please check one)

All of my potentially eligible patients

All patients for which I have checked "Approach" below

None of my patients

Doctor's signature: _____ Date: _____

HOSP NUMBER	Patient Name	Approach	Do Not Approach	Not My Patient	Comment:
ttttt-t	Jane Doe				
xxxxx-x	John Doe				
yyyyy-y	Julie Doe				
zzzzz-z	Paul Doe				

Figure 2. Sample recruitment checklist presented to a primary care provider.

For example, if a patient is eligible for a study and refuses to participate, this information would prevent him or her from showing up on future appointment lists as eligible for that study, but he or she might be eligible for a different study if he or she meets those criteria.

The Regenstrief Data Core has also established the Regenstrief Study Monitoring System. This system uses the above-mentioned master database to upload enrollment dates into patients' RMRS EMRs. This allows clinicians and researchers to identify patients in various studies. The RMRS

contains the dates of enrollment and discharge for each patient. This allows the data managers to establish date ranges for descriptive and outcome clinical data for enrolled patients during their time in the study. It allows them to know when patients have completed a study or have been removed from a study so they may be available for recruitment into other studies.

Electronic Data Collection: Recruitment for Studies of Chronic Conditions

At the onset of a new study, researchers generate criteria for potential study subjects. By working directly with the Regenstrief Data Core and the ResNet administration, an electronic data extraction is performed to generate a list of potential subjects from the RMRS. The data managers are intimately familiar with the RMRS data model and understand how data are represented, how accurate certain parts of the database are, and how they can define disease from these data. Patients are then automatically excluded if they have previously indicated an unwillingness to participate in future trials. ResNet generates a customized list of patients for each physician (Figure 2). The treating physician indicates whether ResNet should approach a patient for enrollment in the study. Eligible patients will be approached by RAs at their next scheduled appointment and taken to a specially designated interview room to discuss their willingness to participate in research. The patient preference and any initial screening information are recorded at this session using portable laptops.

Electronic data management ensures that patient and physician preferences are stored, retrieved, and delivered to minimize inconvenience to both parties and to maximize the efficiency of the RAs.

Electronic Data Collection: Recruitment for Studies of Acute Conditions

The Medical Gopher is one of the oldest and most comprehensive inpatient⁸ and outpatient⁹ physician order-writing workstations extant. It has been used as an intervention itself, resulting in an average of \$900 lower inpatient charges and a one-third reduction in drug-related incident reports.⁸ It has also been used as a platform for presenting information to physicians, including reminders to perform inpatient preventive care and test- and patient-specific information to reduce outpatient testing.¹⁰⁻¹² It is the sole means for writing orders in most of our primary care sites, including the busiest and those used most for research.

ResNet uses the Gopher order entry system to enlist patients. For studies of acute conditions, the Gopher recognizes the term that the physician must enter to indicate the reason for the visit. If the physician enters “back pain” (which can be selected from menus or through partial name matching of a few typed letters) (Figure 3), a window opens on the screen and informs the physician that there is an ongoing study of acute back pain (Figure 4). The most important inclusion and exclusion criteria are displayed. The physician is asked whether a RA can contact the patient for inclusion in the study. If the physician says “yes,” a message is sent via E-mail to

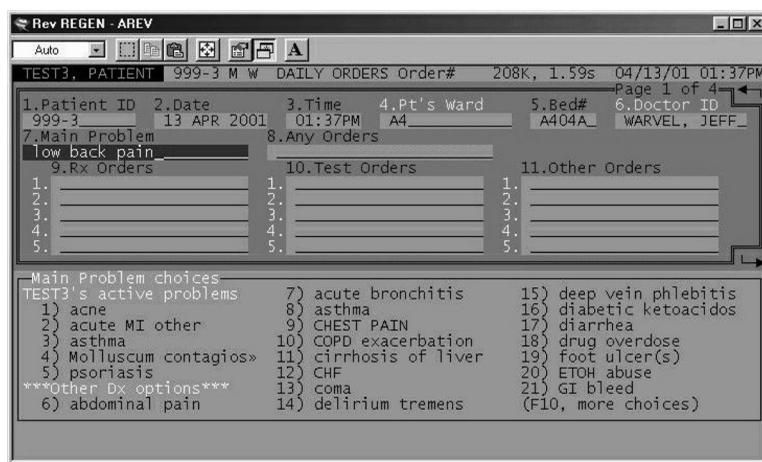


Figure 3. Screenshot of a physician entering a visit reason (Main Problem) in the Regenstrief Gopher order entry system.

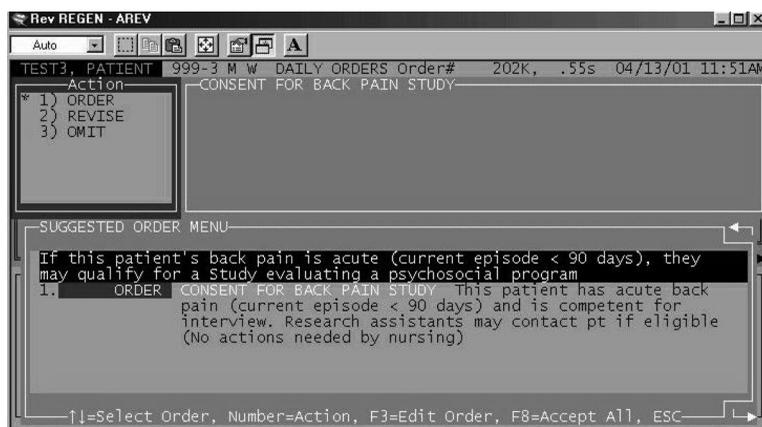


Figure 4. Screenshot of a reminder to the ordering physician that a patient may be eligible for study enrollment as triggered by the visit reason.

a paging Web address and then to a pager carried by the ResNet RA. The patient can then be contacted before leaving the practice and invited to participate in the study.

Electronic Data Collection: Identifying Research Subjects at Follow-up Visits

In April of 2005, the clinics within our PBRN instituted open-access scheduling to reduce the rate of “no-shows.” With open access, patients often would call in only a day or two in advance of their clinic visit. Weekly lists often missed these last-minute registrations.

A redesign of our electronic recruitment system was necessary to capture just-in-time patient appointments. Because our clinics use a registration system compliant with the widely accepted HL7 messaging standard, we were able to make this change readily. As patients register, an electronic HL7 message is generated and sent for storage in a central repository. Using open source software,¹³ we captured these messages to dynamically update our existing patient lists. RAs used these dynamic daily lists of eligible research patients to “triage” themselves to where the most patients will be available for recruitment that day.

After the institution of open access, the number of patients that were absent from the RA lists had increased by 5-fold (Table 1). Institution of our real-time electronic registration capture tool allowed the RAs to once again identify study patients at their follow-up appointments at pre-open-access rates (ie, >99% of study patients identified and approached at follow-up appointments).

Discussion

We described personnel, organizational structures, and technologies that have helped our PBRN function efficiently and successfully. The development of a central repository (which in our case is the RMRS) is of paramount importance because the entire data collection process depends on it. Understanding how data and diseases can be machine represented is critical as well. Embracing standards to represent and transmit clinical data ensures that data collected from any clinic will be consistently represented in our central repository. We described how these processes work in ResNet to demonstrate how using these standards made adaptation to unforeseen changes possible.

Although PBRNs greatly vary with regards to personnel and technical capacity, our successful ap-

Table 1. Missed Patients on Weekly Research Assistant (RA) Lists

Dates	Total Appointments	Number Not on RA Lists	Percentage of Appointments Missing from Schedule
9/01/04 to 4/30/05 (pre-open access)	1538	11	0.72
5/01/05 to 11/17/05 (open access)	3495	132	3.8
11/18/05 to 1/31/06 (intervention)	1727	15	0.86

plications utilized technology available to most other PBRNs. A majority of clinics already register patients electronically.^{14,15} Many of these systems are HL7 compliant. Free open source tools exist for the capture of these messages. And computers are ubiquitous throughout many PBRN clinics.^{16,17} It is worth reiterating this key technical point: HL7 messages are used by most electronic registration systems and represent an opportunity for consistent and real-time data capture. Even in the absence of a fully functional EMR, simply tapping electronic registration messages can allow a centralized researcher with a list of eligible patients to identify these patients in follow-up. And free tools exist to make this a reality.^{13,18}

However, there exist a number of barriers to IT adoption in PBRNs. We are fortunate to have an extensively studied EMR system and computerized provider order entry ingrained into the daily practice of our clinics. Few other institutions have such an IT culture, although interest in adoption appears high.¹⁵

Rather, much research data are stored on stand-alone computer databases. Many commercial EMR systems are not set up for the client to extract data on their own. In this instance, it may be more possible to aggregate data streams (from labs, hospitals, clinics, insurance providers, etc) into a data repository before storage in the EMR system. And personnel, technical, and financial resources may be in short supply. Estimates of infrastructure costs (personnel, administrative, hardware/software) range from \$69,700 to \$287,600 per year for a moderately complex network.¹⁹

Despite these barriers, IT can and will be a powerful tool to dramatically increase the efficiency and success of PBRN research. All evidence points to an ever-expanding role for IT in the delivery of health care.²⁰ Clinicians and researchers play a critical role in shaping this adoption process. Understanding the principles and key concepts for planning and implementation will ensure the greatest chance of success.

We thank Brenda Hudson, Dede Willis, Evgenia Teal, Mike Plue, Faye Smith, and Jane French for invaluable contributions to ResNet.

References

- Breitfeld PP, Weisburd M, Overhage JM, Sledge G, Jr, Tierney WM. Pilot study of a point-of-use decision support tool for cancer clinical trial eligibility. *J Am Med Inform Assoc* 1999;6:466–77.
- McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectability of man. *N Engl J Med*. 1976;295:1351–5.
- Huff SM, Rocha RA, McDonald CJ, et al. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. *J Am Med Inform Assoc* 1998;5:276–92.
- SNOMED.org [homepage on the Internet]. Northfield (IL): SNOMED International; c2006 [cited 20 Sept 2006]. Available from: <http://www.snomed.org>.
- HL7.org [homepage on the Internet]. Ann Arbor (MI): Health Level Seven, Inc.; c1997–2006 [cited 20 Sept 2006]. Available from: <http://www.hl7.org>.
- Green LA, White LL, Barry HC, Nease DE, Hudson BL. Infrastructure requirements for practice-based research networks. *Ann Fam Med* 2005;3 Suppl 1:S5–11.
- McDonald CJ, Overhage JM, Tierney WM, et al. The Regenstrief Medical Record System: a quarter century experience. *Int J Med Inform* 1999;54:225–53.
- Tierney WM, Miller ME, Overhage JM, McDonald CJ. Physician inpatient order writing on microcomputer workstations: effects on resource utilization. *JAMA* 1993;269:379–83.
- McDonald CJ, Tierney WM. The Medical Gopher—a microcomputer system to help find, organize and decide about patient data. *West J Med* 1986;145:823–9.
- Tierney WM, McDonald CJ, Martin DK, Hui SL, Rogers MP. Computerized display of past test results: effects on outpatient testing. *Ann Intern Med* 1987;107:569–74.
- Tierney WM, McDonald CJ, Hui SL, Martin DK. Computer predictions of abnormal test results: effects on outpatient testing. *JAMA* 1988;259:1194–8.
- Tierney WM, Miller ME, McDonald CJ. The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. *N Engl J Med* 1990;333:1499–504.
- HL7 Comm. Mike Litherland [cited 20 Sept 2006]. Available from: <http://nule.org>.
- Ariza AJ, Binns HJ, Christoffel KK. Evaluating computer capabilities in a primary care practice-based research network. *Ann Fam Med* 2004;2:418–20.
- Andrews JE, Pearce KA, Sydney C, Ireson C, Love M. Current state of information technology use in a US primary care practice-based research network. *Inform Prim Care* 2004;12:11–18.

16. Pace WD, Staton EW. Electronic data collection options for practice based research networks. *Ann Fam Med* 2005;3 Suppl 1:S21-9.
17. Nutting PA. Practice-based research networks: building the infrastructure of primary care research. *J Fam Pract* 1996;42:199-203.
18. HL7 Tools, Utilities and Resources [homepage on the Internet]. Avalon (Australia): HL7 Systems & Services. c2001-06 [updated Dec 2006; cited 20 Sept 2006]. Available from: <http://www.hl7.org.au/HL7-Tools.htm>.
19. Green LA, White LL, Barry HC, Nease DE, Hudson BL. Infrastructure requirements for practice based research networks. *Ann Fam Med* 2005;3 Suppl 1:S5-11.
20. Ortiz E, Clancy CM. Use of information technology to improve the quality of health care in the United States. *Health Serv Res* 2003;38:11-22.