A Model for the Electronic Support of Practice-Based Research Networks

Kevin A. Peterson, MD, MPH[1]

Brendan C. Delaney, MD, FRCP, MRCGP[2]

Theodoros N. Arvanitis, RT, DPhil[3]

Adel Taweel, PhD[2]

Elisabeth A. Sandberg, MS, MIM, MBA[1]

Stuart Speedie, PhD, MED[1]

F. D. Richard Hobbs, FMedSci, FRCGP[4]

[1]University of Minnesota, Minneapolis, Minnesota

[2]King's College, London, United Kingdom

[3]Birmingham University, Edgbaston, United Kingdom

[4]Oxford University, Oxford, United Kingdom

ABSTRACT

PURPOSE The principal goal of the electronic Primary Care Research Network (ePCRN) is to enable the development of an electronic infrastructure to support clinical research activities in primary care practice-based research networks (PBRNs). We describe the model that the ePCRN developed to enhance the growth and to expand the reach of PBRN research.

METHODS Use cases and activity diagrams were developed from interviews with key informants from 11 PBRNs from the United States and United Kingdom. Discrete functions were identified and aggregated into logical components. Interaction diagrams were created, and an overall composite diagram was constructed describing the proposed software behavior. Software for each component was written and aggregated, and the resulting prototype application was pilot tested for feasibility. A practical model was then created by separating application activities into distinct software packages based on existing PBRN business rules, hardware requirements, network requirements, and security concerns.

RESULTS We present an information architecture that provides for essential interactions, activities, data flows, and structural elements necessary for providing support for PBRN translational research activities. The model describes research information exchange between investigators and clusters of independent data sites supported by a contracted research director. The model was designed to support recruitment for clinical trials, collection of aggregated anonymous data, and retrieval of identifiable data from previously consented patients across hundreds of practices.

CONCLUSIONS The proposed model advances our understanding of the fundamental roles and activities of PBRNs and defines the information exchange commonly used by PBRNs to successfully engage community health care clinicians in translational research activities. By describing the network architecture in a language familiar to that used by software developers, the model provides an important foundation for the development of electronic support for essential PBRN research activities.

Ann Fam Med 2012;10:560-567. doi:10.1370/afm.1434.

Conflicts of interest: authors report none.

CORRESPONDING AUTHOR

Kevin Peterson, MD, MPH
717 Delaware St SE, Ste 425
Minneapolis, MN 55414
peter223@umn.edu

INTRODUCTION

Networks of primary care practices working cooperatively to address clinical research questions now provide promising new laboratories for discovery.[1,2] Bound together by a shared commitment to improving health care through systematic inquiry, practice-based research networks (PBRNs) provide coordination for thousands of practices across the United States in order to investigate health care events that occur in community settings.[3] The ability of PBRNs to involve real-world practices in clinical research extends opportunities for engagement with previously unavailable populations and promotes rapid adoption of research findings by community practices.[4]

Although the introduction of electronic health records in primary care increases the potential for PBRNs to electronically involve practices in

research, PBRNs have faced challenges in successfully adopting currently available health information technology.[5] Research designs and methodologies used by PBRNs differ in important ways from methods typically used by academic research enterprises or large managed care organizations with centralized research authorities. Electronic systems supporting data collection within a PBRN must account for wide variation in local resources and organizational workflows that prioritize clinical care delivery. Health information technology must support standardization of distributed data across heterogeneous data sources, accommodate a wide variety of study designs, and integrate PBRN roles. Data sharing must comply with existing business rules from independent practices and medical systems. Improved technology is necessary for the rapid translation of new research into practice and ultimately for rapid health care improvement integral to the promise of PBRNs.[4] Until PBRN roles and activities are more accessible to software developers through standardized models, technology is unlikely to provide robust support.

The electronic Primary Care Research Network (ePCRN) was funded in 2005 by Re-engineering the Clinical Research Enterprise, a Roadmap Initiative of the National Institutes of Health, to determine feasibility for a National Electronic Clinical Trials and Research network (NECTAR).[6] The principal goal was to enable the development of an electronic infrastructure to support clinical research activities in primary care PBRNs. We describe a model that was developed by the ePCRN to enhance the growth and to expand the reach of PBRN research.

METHODS

Use Cases

We collected study designs and research processes from studies undertaken by 11 PBRNs from the United States and United Kingdom (listed at the end of this article). Studies were evaluated through key informant interviews and group discussions that captured the actors, processes, dependencies, and data types used in practice-based research. Initial qualitative analysis identified 3 broad categories of clinical research involving patients with distinct dependencies and data flows: clinical trial recruitment, collection of aggregated anonymous data, and collection of data from consented patients.[7] We defined 3 principal use cases from these categories to provide support for most PBRN activities: (1) screening and recruitment of selected individuals from community sites for a clinical trial, (2) rapid collection of aggregated anonymous data from a defined population across multiple community sites under a Health Insurance Portability and

Accountability Act waiver, and (3) retrieval of identified data from a previously consented patient.

Architecture Modeling

We modeled roles and activities from each use case using Unified Modeling Language. Each use case was broken down into specific activities, and activity diagrams were created on a visual modeling platform (Enterprise Architect, Sparx, version 9.1). Proposed models were evaluated, defined, and revised by researchers, PBRN staff, and community clinicians to ensure that all existing studies were supported by one of the use cases. The use cases were used to drive development and testing of functional components, architecture, and governance.

We constrained modeling whenever possible to be consistent with standards from existing research object models from the Clinical Data Interchange Standards Consortium (CDISC), Health Level 7 (HL7) Regulated Clinical Research Information Model (RCRIM), and the Biomedical Research Integrated Domain Group (BRIDG) Model.[8] Activities were organized into so-called swimlanes corresponding to different functional roles (actors). Similar activities in different use cases were identified, refined, and shared. Clusters of activities requiring similar support capabilities were joined into discrete functional components.

Actors

We defined 3 organizational roles: practices, PBRNs, and research organizations. Each organization has individual actors. The overall data flow between organizations is shown in Figure 1, and activities are summarized below. Although organizations, roles, and activities are described in terms familiar to PBRNs, the terms can be applied in many ways. For example, in PBRNs that have individual clinicians as members instead of clinics, practice may refer to an individual clinician or group of clinicians. Using this interpretation, multiple practices could exist within a single clinic. In a managed care organization, a practice may represent a hospital or specific database requiring independent or technical oversight. In this case, the PBRN role could be assigned to an affiliated research institute. In an academic center, practices may represent individual clinics and the PBRN Director role could be assigned to a Clinical Translational Science Institute with multidisciplinary oversight.

Practices

A practice is a data owner. Practices can be a member of 1 or more PBRNs. The practice director is the data steward. Clinicians and staff participate in clinical research, coordinate patient consent, and manage patient involvement in specific studies. Practice activities include both research and local data analysis. Research activities

include registering with a PBRN, reviewing information on potential studies, reviewing requests from researchers, obtaining consent for screening, and data sharing of aggregate or individual data. Local data analysis activities include data queries for population health management, quality improvement, disease registry functions, and practice and clinician performance reports.

PBRNs

A PBRN is a collection of practices that have agreed to jointly conduct 1 or more research studies. The PBRN is administered by a network director supported by staff. PBRNs enhance the value of research participation for practices, ensure compliance with regulatory requirements, and promote the maintenance of high-quality data in practices. PBRNs draw on research experience from practices to assist in selection of studies, better understand potential impacts on clinical workflows, and identify hidden costs of participation.

Research Organizations

The research organization supports researchers and facilitates access to PBRN resources. Researchers may propose a new study, develop specific data queries, initiate research requests, select PBRNs, and negotiate details of practice participation. Study data are returned to the research organization and managed by the researcher.

Data Governance

Data are owned by practices. We included the following as important data governance issues: (1) practices

maintain autonomy, independence, and complete control over their data; (2) the practice director is responsible for ensuring appropriate sharing of practice data; (3) every study in a practice must comply with local data privacy requirements and accommodate local concerns and community sensitivities about data sharing; (4) although PBRNs have access to clinician names and summary data (counts) from each practice, they do not have access to protected health information unless released by the practice with documented patient consent; (5) data queries are locally recorded to comply with disclosure requirements; (6) notification plans are in place in case of a potential data privacy breach; and (7) data are reviewed by the practice director before release to the PBRN.
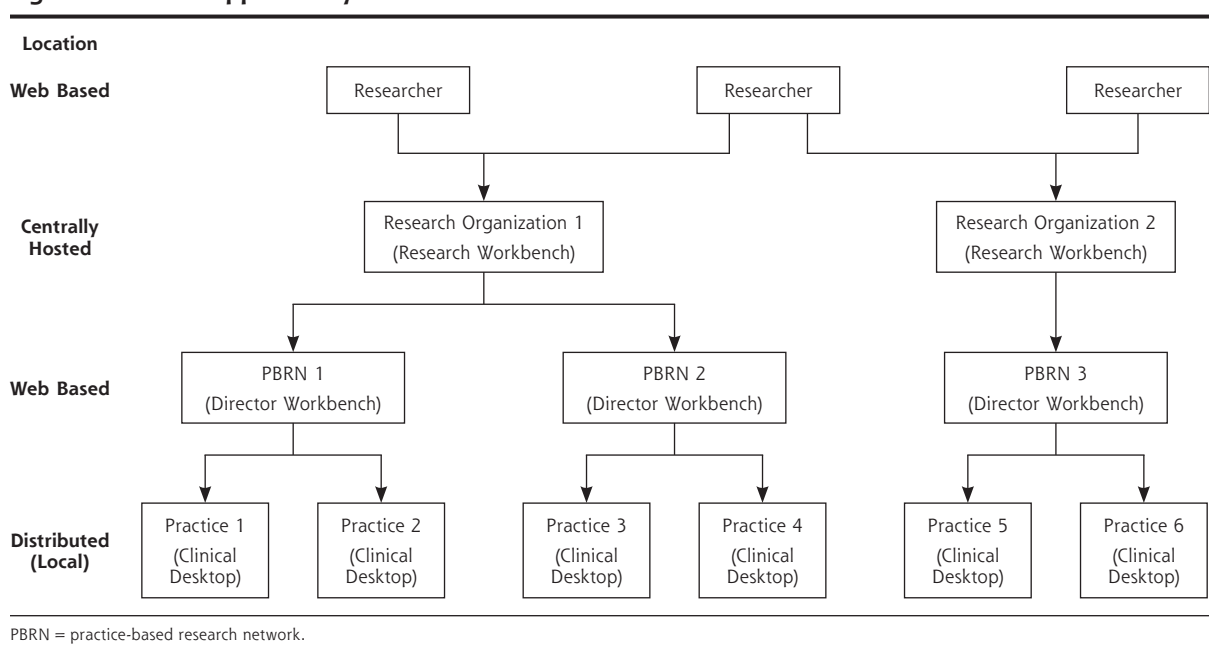
The PBRN network director provides stewardship of data after release by the practice but before release to research organizations. Practices and clinicians are anonymized. Data are reviewed by the PBRN before release to the research organization providing a double-check for regulatory compliance. Aggregate data are checked to ensure correlations do not compromise anonymity.

Once data are released by the PBRN to a research organization, the researcher becomes the data steward. Data are subject to protection of human subjects restrictions and any negotiated data use agreements.

Business Rules

We included the following business rules for practices: (1) patient participation is defined by existing patient

Figure 1. Actors supported by the Web architecture.

Location			
Web Based	Researcher	Researcher	Researcher
Centrally Hosted	Research Organization 1 (Research Workbench)		Research Organization 2 (Research Workbench)
Web Based	PBRN 1 (Director Workbench)	PBRN 2 (Director Workbench)	PBRN 3 (Director Workbench)
Distributed (Local)	Practice 1 (Clinical Desktop) Practice 2 (Clinical Desktop)	Practice 3 (Clinical Desktop) Practice 4 (Clinical Desktop)	Practice 5 (Clinical Desktop) Practice 6 (Clinical Desktop)

PBRN = practice-based research network.

agreements in the practice; (2) patients can opt out and not contribute data to research, but all patients are included in local quality improvement work done by the practice; and (3) participation in any PBRN study by a practice is voluntary and study specific.

We also included several PBRN business rules: (1) PBRNs assist geographically distributed practices to participate in research; (2) one or many sites may participate in any study; (3) practices can be voluntary members of 1 or more PBRNs; (4) PBRNs assist practices in providing high-quality data the through training and resource support; (5) PBRNs provide research expertise to assist practices in stewardship of data; and (6) PBRNs distribute studies to selected practices to promote efficiency and comply with individual practice restrictions.

Business Model

PBRNs negotiate with researchers for fair reimbursement to practices for the additional costs of research participation. PBRNs provide support to practices to encourage participation in research and balance the needs of researchers and practices. Researchers select among PBRNs to minimize cost and maximize data quality. Practices join PBRNs to maximize the value of research participation. As practices differ in their perception of value, factors driving practice participation vary with the availability of research resources, size of the practice data set, engagement of clinicians, and perceived value to patients. In many practices the intrinsic value of the data is small and would not offset the disruption of workflow caused by introducing research. To enhance the value of participation in these cases, PBRNs may provide additional incentives to practices including quality improvement tools, disease registries, and data analysis tools leveraged by sharing the cost across many practices.

Web Architecture

We separated functional components defined under architecture modeling into packages reflecting hardware requirements, actors, data governance, and PBRN business rules. To decouple dependencies between the packages, web service communication was designed to be independent of internal data structure. Packages identified as "services" provide a reference service that could be used by multiple PBRNs. Any package can be interchanged with a customized application using similar web services. The ability to replicate, improve, and replace packages enhances scalability and supports a customized distributed model.

RESULTS

Table 1 shows the functional components of PBRN research activities identified through architectural modeling of existing PBRN research. Prototype software for these functional components was created, and a test of the prototype in 100 practices has been reported.[9] Several research activities previously not described by existing research object models were identified and published.[10]

Figure 2 presents an overall composite diagram of the ePCRN model. The 7 structural components (applications) in bold text show essential interactions,

Table 1. Functional Components Contained in Each Package

Functional Component	Activities
Study workbench	Protocol capture
	Auditing
	Eligibility capture
	Participant counting
	Participant flagging
	Distributed querying
	Deidentified data importer
	Deidentified DB
	Ontology service
	Ontology support interface
	Identified data importer
	Identified DB
	Recruitment tracker
	Data collector
Client-SOAP interface engine	Security layer
Clinic desktop	Data migration and import
	Clinical performance DB clinical performance reports
	Deidentified DB
	Identified DB
	Flagging tool
	Patient opt out tool
	Recruitment helper
	Data review and transfer
Director workbench	Role assignment
	Request capture
	Query review
	Query assignment
	Approval
	Audit
Study storage services	Data collection
	Data review and import
	Results database
	Security data
Ontology services	Ontology support interface
	Ontology DB
	Update service
Study information services	Study information DB
Administrative services	Clinic registry
	Certificate authority
Install manager	Register clinic
	Generate host certificate
	Start/stop service

DB = database; SOAP = Simple Object Access Protocol.

Figure 2. Composite diagram of proposed major components for PBRN support model.

PBRN = practice-based research network; IRB = institutional review board; EHR = electronic health record; PHR = personal health record.
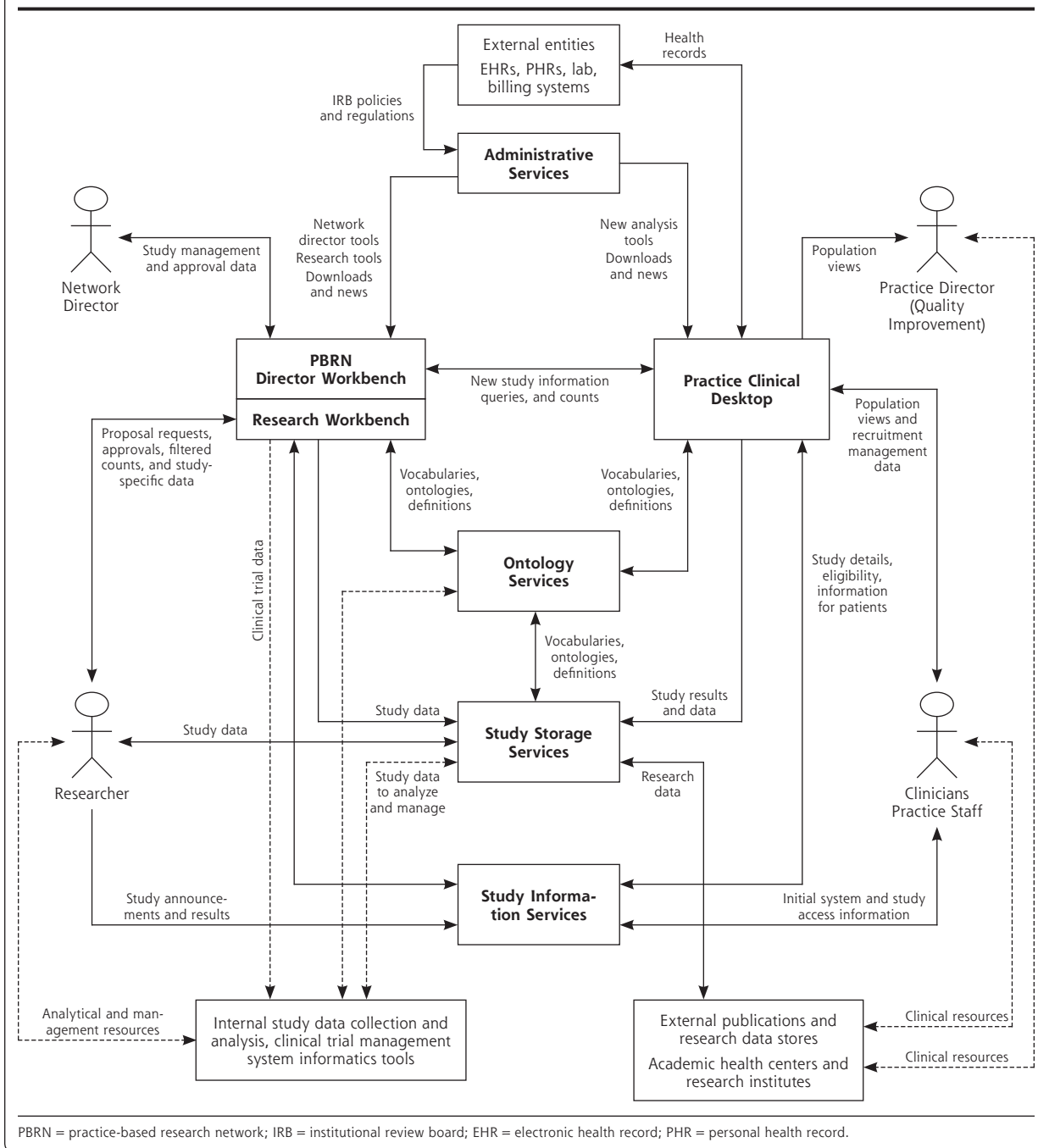
activities, and data flows: administrative services, director workbench, research workbench, practice clinical desktop, ontology services, study storage services, and study information services. Details of these components are described below. In a general sense, the model describes an electronic infrastructure for secure information exchange between researchers and clusters of heterogeneous and independent practices coordinated through PBRNs. The model supports all use

cases, data governance, and business rules described for PBRNs. The overall function of individual components is summarized below, and a scenario for a single use case is described.

Components

The practice clinical desktop provides a standardized framework for interaction at the practice and supports all identified research and local data analysis require-

ments identified for practices. Owned by the practice, the application can be installed at the practice or hosted on the web.

The clinical desktop provides the practice director with a flexible interface for quality improvement and local queries of identifiable data using a highly scalable star schema. Functions include support for migration and loading of the data set. Although unmapped data can be added to the data set, the clinical desktop interacts with ontology services to facilitate data mapping and maintenance. Data migration tools will be described further in a separate publication. Although data mapping is not necessary for local queries, mapped data sets enable external queries and provide better support for PBRN and practice business models.

The clinical desktop controls participation in PBRN data sharing. Data queries can be limited by PBRN-defined roles, restricted by data type, or not allowed. The clinical desktop facilitates consent for screening within the practice. Aggregate data are stripped of identifiers.

The PBRN director workbench supports data governance and business rules described for PBRNs. Hosting may be determined by the PBRN; however, the prototype supports hosting at the research organization. The application provides for assignment of administrative roles to staff, facilitates negotiation with researchers, checks for presence of targeted data by practice, tracks approval of research requests, tracks data collection, and reviews aggregated data before delivery to the research organization.

The director workbench interacts with the administrative services package, which registers and tracks practices, assigns research queries, and adjusts distributed queries to match version differences in the clinical desktop package.

The research workbench is hosted by the research organization and facilitates the creation and initiation of research studies.

The research workbench passes study information to study information services, where it is accessible by PBRNs and practices. Documents uploaded to study information services can be provided to practices or patients to facilitate recruitment. The research workbench supports aggregate counts of data to assist in ascertaining availability of data and tracks data collection.

Data are returned to study storage services, where they are available to authorized researchers in a relational database.

Scenario

A typical scenario for identifying a patient for a clinical trial helps to explain how the information architecture

in Figure 2 works. The researcher logs onto a research workbench and defines a study using World Health Organization standards. Study documents are uploaded to study information services, where they are available to the other actors. The researcher submits a count to establish the presence of targeted data or to measure the impact of specific inclusion and exclusion criteria on the available population. Creation of the query is supported by ontology services that provide metadata tags to enhance capture across heterogeneous data sets. The queries are translated into multiple coding schemes and arranged by administrative services to coincide with the clinical desktop version in the selected practices before being securely distributed. If the PBRN and the practice have decided to automatically allow anonymous counts, the summary count is aggregated by the PBRN and returned to the research workbench. The researcher selects a PBRN and requests approval for engaging member practices in a study.

The network director reviews research requests ensuring that requests meet local regulatory requirements, assesses potential impacts on workflow and local resources, negotiates costs, and directs the study to the most suitable practices. Practices may review study requests and decide whether or not to participate. The practice is provided with the list of eligible patients in the practice and access to study details from study information services. Patients may be contacted and consented by the practice, screened, or released for screening by the researcher in compliance with local human subjects protection approvals. After consent, contact information is released to the designated study storage service. Data may be held at the study storage services until reviewed by the network director, providing a double-check on regulatory compliance.

DISCUSSION

PBRNs support engagement and collaboration with community clinicians that is essential for advancing discovery and disseminating new research findings. The Agency for Healthcare Research and Quality PBRN Resource Center has registered more than 150 PBRNs in the United States with 67,000 clinicians from 16,500 clinical practices.[11] As electronic health records become widely adopted, research networks that span multiple distributed data sources can provide powerful tools for clinical research. The proposed model provides an important foundation for the development of software supporting PBRN research activities.

The model identifies 7 independent applications that interact over the Internet using standard web services. Components support local customization and enhancement. Although 3 principal use cases are addressed,

many additional use cases are possible. Support for quasi-experimental or multimethod studies could be provided by combining existing functions, while direct patient involvement at the clinical desktop could be envisioned to support community-based participatory research and provide a more systematic approach to qualitative research in community settings. The model scales to a size large enough to support every PBRN and every primary care practice in the United States. The model complements the function of clinical trial management software but does not replace it.

Although creation of a compatible electronic infrastructure is a considerable hurdle for the electronic support of PBRNs, other important barriers still exist. Despite ownership, migration of data into the clinical desktop data set is an obstacle for practices and has delayed US implementation. Although data migration tools provide substantial support, collection and standardization of high-quality data remains a focus of ongoing work. Implementation of the ePCRN model in the European Union with several large clinical trials has proceeded quickly under the TRANSFoRm project "to develop a 'rapid learning healthcare system' that can improve both patient safety and the conduct and volume of clinical research in Europe."[12]

The potential for introducing a fast and efficient infrastructure to facilitate PBRN research offers the possibility of rapid advances in a wide variety of areas including comparative effectiveness research, patient safety, event monitoring for drugs and devices, and clinical trials. Engagement of community practices and clinical data without compromising the privacy of individuals, clinicians, or practices provides a leap forward in available tools for evaluating the effectiveness of the health care system.

The proposed model is limited in a number of ways. The model is designed to reflect electronic support and does not account for other types of community engagement. The model may not provide robust support to PBRNs or other research networks that have already established different research workflows or have specific needs not typical of PBRNs. The model was developed in the United States and United Kingdom, and may not appropriately reflect research activities in other countries. Finally, although the separation of research activities into specific software packages is intended to reflect existing PBRN business rules, different interpretations about which research activities should be assigned to specific actors could be made by experienced research institutions and would lead to a different model for the proposed service interactions.

Other electronic research architectures supporting distributed research queries have been proposed. I2B2 Clinical Chart is an Eclipse-based solution that sup-

ports distributed queries of anonymous data sets using the SHRINE tool.[13] Although some ePCRN functionality, such as deidentified aggregate counts, are also found in the I2B2 SHRINE environment, SHRINE does not support many activities necessary for successful research within PBRNs including the provision for local data oversight on a study-by-study basis. Potential methods of integrating ePCRN software within an I2B2 framework have been presented elsewhere.[14] Some health systems and managed care organizations have developed tools that accomplish similar functions; however, they have not addressed specific PBRN requirements and roles.[15]

Differences between the US and UK health systems have required customization of specific modules to accommodate distinct national regulatory requirements. Although substantial differences in implementation strategy and business rules exist, cooperation between the United States and the European Union has ensured that the model has the flexibility to accommodate international collaboration between PBRNs. The ePCRN Consortium, formed in 2010 by the University of Minnesota, King's College London, and Birmingham University (England), provides an open forum to promote continued development of electronic support for PBRNs and to ensure compatibility between components. Software developed in the Consortium is freely available in the United States under the ePCRN Open Source License (http://www.epcrn.net).

To read or post commentaries in response to this article, see it online at http://www.annfammed.org/content/10/6/560.

Participating networks: Alan Adelman, MD, MS, Penn State Ambulatory Research Network; Deborah Allen, MD, Indiana Family Practice Research Network; Walter Calmbach, MD, South Texas Ambulatory Research Network; Myra Crawford, PhD, MPH, Alabama Practice-Based Research Network; Pat Fontaine, MD, Minnesota Academy of Family Physicians Practice-Based Research Network; Chet Fox, MD, Upstate New York Practice-Based Research Network; Lyndee Knox, MD, Los

Angeles Practice-Based Research Network; Zsolt Nagykaldi, PhD, Oklahoma Physicians Resource/Research Network; Wilson Pace, MD, AAFP National Research Network and State Networks of Colorado Ambulatory Practices and Partners; John Ryan, DrPH, South Florida Primary Care Practice-Based Research Network; Brendan C. Delaney, Midlands Research Practices Consortium (MidReC), Birmingham, England.

References

1. Institute of Medicine. *Primary Care: America's Health in a New Era.* Washington, DC: National Academy Press; 1996.

2. Institute of Medicine. *Report from the Committee on the Future of Primary Care.* Washington, DC: IOM; 1998.

3. Westfall JM, Mold J, Fagnan L. Practice-based research—"Blue Highways" on the NIH Roadmap. *JAMA.* 2007;297(4):403-406.

4. Lindbloom EJ, Ewigman BG, Hickner JM. Practice-based research networks: the laboratories of primary care research. *Med Care.* 2004;42(4 Suppl):III45-III49.

5. Nagykaldi Z, Mold JW. The role of health information technology in the translation of research into practice: an Oklahoma Physicians Resource/Research Network (OKPRN) study. *J Am Board Fam Med.* 2007;20(2):188-195.

6. Peterson KA, Fontaine P, Speedie S. The electronic Primary Care Research Network (ePCRN): a new era in practice-based research. *J Am Board Fam Med.* 2006;19(1):93-97.

7. Pace WD, Cifuentes M, Valuck RJ, Staton EW, Brandt EC, West DR. An electronic practice-based network for observational comparative effectiveness research. *Ann Intern Med.* 2009;151(5):338-340.

8. Fridsma DB, Evans J, Hastak S, Mead CN. The BRIDG project: a technical report. *J Am Med Inform Assoc.* 2008;15(2):130-137.

9. Fontaine P, Mendenhall TJ, Peterson K, Speedie SM. The "Measuring Outcomes of Clinical Connectivity" (MOCC) trial: investigating data entry errors in the electronic Primary Care Research Network (ePCRN). *J Am Board Fam Med.* 2007;20(2):151-159.

10. Speedie SM, Taweel A, Sim I, Arvanitis TN, Delaney B, Peterson KA. The Primary Care Research Object Model (PCROM): a computable information model for practice-based primary care research. *J Am Med Inform Assoc.* 2008;15(5):661-670.

11. Peterson K, Durako S, eds. *The AHRQ PBRN Resource Center Update. AHRQ PBRN 2011 National Conference Proceedings.* Washington, DC: Association for Healthcare Research and Quality, Department of Health and Human Services; 2011.

12. Institute of Medicine. *Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care: Workshop Series Summary.* Washington, DC: National Academies Press; 2011:197-204.

13. Weber GM, Murphy SN, McMurry AJ, et al. The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. *J Am Med Inform Assoc.* 2009;16(5):624-630.

14. Abend A, Peterson K, Anderson N, Mandel A. ePCRN and i2b2— An Architecture for Integrating Primary Care Data into a Research Data Platform. Proceedings from AMIA 2010 Summit on Clinical Research Informatics. March 13, 2010, San Francisco, CA.

15. Holmes JH, Brown J, Hennessy S, et al. Developing a distributed research network to conduct population-based studies and safety surveillance. *AMIA Annu Symp Proc.* 2008;Nov 6:973.