



OPEN ACCESS



Open Access  
Scan to access more  
free content

# CAPriCORN: Chicago Area Patient-Centered Outcomes Research Network

Abel N Kho,<sup>1</sup> Denise M Hynes,<sup>2,3</sup> Satyender Goel,<sup>1</sup> Anthony E Solomonides,<sup>4</sup> Ron Price,<sup>5</sup> Bala Hota,<sup>6</sup> Shannon A Sims,<sup>6</sup> Neil Bahroos,<sup>7</sup> Francisco Angulo,<sup>8</sup> William E Trick,<sup>8</sup> Elizabeth Tarlov,<sup>9</sup> Fred D Rachman,<sup>10</sup> Andrew Hamilton,<sup>10</sup> Erin O Kaleba,<sup>10</sup> Sameer Badlani,<sup>11</sup> Samuel L Volchenbom,<sup>12</sup> Jonathan C Silverstein,<sup>4</sup> Jonathan N Tobin,<sup>13</sup> Michael A Schwartz,<sup>3</sup> David Levine,<sup>14</sup> John B Wong,<sup>15</sup> Richard H Kennedy,<sup>5</sup> Jerry A Krishnan,<sup>2,7</sup> David O Meltzer,<sup>11</sup> John M Collins,<sup>16</sup> Terry Mazany,<sup>17</sup> for the CAPriCORN Team

For numbered affiliations see end of article.

## Correspondence to

Dr Abel N Kho, Division of General Internal Medicine, Northwestern University, 750 N Lake Shore Drive, 10th Floor, Chicago, IL 60611, USA; Abel.kho@nmff.org

Received 25 March 2014

Revised 4 April 2014

Accepted 8 April 2014

Published Online First

12 May 2014



- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002864>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002747>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002758>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002764>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002744>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002746>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002743>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002827>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002740>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002727>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002751>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002745>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002759>
- ▶ <http://dx.doi.org/10.1136/amiajnl-2014-002756>



CrossMark

**To cite:** Kho AN, Hynes DM, Goel S, et al. *J Am Med Inform Assoc* 2014;**21**:607–611.

## ABSTRACT

The Chicago Area Patient-Centered Outcomes Research Network (CAPriCORN) represents an unprecedented collaboration across diverse healthcare institutions including private, county, and state hospitals and health systems, a consortium of Federally Qualified Health Centers, and two Department of Veterans Affairs hospitals. CAPriCORN builds on the strengths of our institutions to develop a cross-cutting infrastructure for sustainable and patient-centered comparative effectiveness research in Chicago. Unique aspects include collaboration with the University HealthSystem Consortium to aggregate data across sites, a centralized communication center to integrate patient recruitment with the data infrastructure, and a centralized institutional review board to ensure a strong and efficient human subject protection program. With coordination by the Chicago Community Trust and the Illinois Medical District Commission, CAPriCORN will model how healthcare institutions can overcome barriers of data integration, marketplace competition, and care fragmentation to develop, test, and implement strategies to improve care for diverse populations and reduce health disparities.

## INTRODUCTION

We established the Chicago Area Patient-Centered Outcomes Research Network (CAPriCORN), with coordination by the Chicago Community Trust (CCT) and the Illinois Medical District Commission (IMDC) and a participatory and nimble governance structure across a diverse group of healthcare institutions including private, county, and state hospitals and health systems, a consortium of Federally Qualified Health Centers (FQHCs), and two Department of Veterans Affairs Hospitals. CAPriCORN seeks to model how healthcare institutions in complex urban settings can overcome the barriers of competition, care fragmentation, and limited resources to develop, test, and implement strategies to improve care for diverse populations and reduce health disparities. The diverse healthcare settings and populations within the CAPriCORN Clinical Data Research Network (CDRN) also serve as a natural laboratory in which we can examine and address the heterogeneity of treatment effects and contribute to the national Patient-Centered Clinical Research Network (PCORnet).

## PARTICIPATING HEALTH SYSTEMS

CAPriCORN brings together an unprecedented Chicago-wide collaboration between 11 diverse healthcare institutions and multiple partner institutions (table 1). Healthcare institutions include: academic medical centers (Loyola University Health System (LUHS), Northwestern Medicine (NM), NorthShore University HealthSystem (NS), Rush University Medical Center (RU), University of Chicago (UC), and the University of Illinois Hospital and Health Sciences System (UI)); Cook County Health and Hospital System (CCHHS); the Alliance of Chicago's FQHCs (Alliance); two local Department of Veteran's Affairs Hospitals and clinics (HinesVAH and Jesse Brown VA (JBVAMC)); and leading pediatric hospitals (Lurie Children's Hospital, Children's Hospital of University of Illinois, and University of Chicago Medicine Comer Children's Hospital). Together, CAPriCORN healthcare institutions provide primary healthcare to over one million patients who mirror the great socioeconomic and racial diversity of our region. Insurance coverage varies from over 70% uninsured at CCHHS to over 70% privately insured at others. All CAPriCORN institutions serve patients with either Medicare or Medicaid coverage. Collectively, our institutions include over 2000 primary care providers (inclusive of Doctors of Medicine (MD), Doctors of Osteopathic Medicine (DO), Nurse Practitioners (NPs), Physician Assistants (PAs)) and over 5000 specialty care providers providing care at 18 distinct hospitals and 463 distinct clinic sites. This diversity of care sites (spanning inpatient, outpatient, and emergency care) creates a natural laboratory for comparative effectiveness research (CER) and includes individual institutions covering specific patient populations who are primarily uninsured, minority, military veterans, Spanish speaking, homeless, or members of other historically vulnerable populations (eg, Lesbian, Gay, Bisexual, and Transgender (LGBT)).

## INFORMATION SYSTEMS AND DATA STANDARDS

CAPriCORN institutions represent a broad range of information systems. All have electronic health record (EHR) systems with components supported by major vendor systems, including Epic (NM, NS,

**Table 1** Data systems and care sites of participating CAPriCORN institutions

Institution name	EHR used	Data warehouse	Owned/operated	
			Hospitals	Clinic sites
Alliance of Chicago	GE Centricity	Parallel Data Warehouse	0	60
Cook County Health and Hospital System	Cerner	Research Data Warehouse	2	14
Loyola University Health System	Epic, Sunquest, GE Centricity	Epic—Clarity Repository, CRDB	2	22
NorthShore University HealthSystem	Epic	Oracle-based EDW	4	134
Northwestern Medicine	Cerner, Epic, eClinicalworks	EDW	2	28
Rush University Medical Center	Epic	EDW	2	75
University of Chicago	Epic, GE Centricity	Clinical Research Data Warehouse	2	2
University of Illinois Hospital and Health Sciences System	Cerner, Epic, Sunquest/MYSYS	Custom Research Data Warehouse	1	37
Jesse Brown VA Medical Center	Veterans Health Information Systems and Technology Architecture (VistA)	CDW	1	4
Edward Hines Jr VA Hospital	VistA	CDW	1	6

CAPriCORN, Chicago Area Patient-Centered Outcomes Research Network; CDW, Corporate Data Warehouse; CRDB, Clinical Research Database; EDW, Enterprise Data Warehouse; EHR, electronic health record; VA, Veterans Affairs.

RU, LUHS, UI), Cerner (UI, CCHHS), GE Centricity (Alliance), and VistA (JBVAMC, HinesVAH). The information systems at all CAPriCORN institutions include a relational data warehouse with structured reporting functionality. With the exception of the Alliance, whose warehouse is limited to outpatient care data from the 11 FQHCs they serve, the CAPriCORN institution warehouses include extensive administrative and clinical data, spanning both the inpatient and outpatient setting.

All institutions currently use International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) coding for diagnoses<sup>1</sup> and Current Procedural Terminology (CPT) codes for procedures.<sup>2</sup> Some institutions use Logical Observation Identifiers Names and Codes (LOINC) codes for laboratory tests, and almost all others have made significant progress towards LOINC implementation.<sup>3</sup> Image data and progress notes are electronic, but the format varies across sites, and some image data are maintained in separate systems at each facility. All institutions have either achieved, or are progressing toward, meaningful use compliance in 2014.

### SUCCESSFUL HISTORY OF DATA EXCHANGE

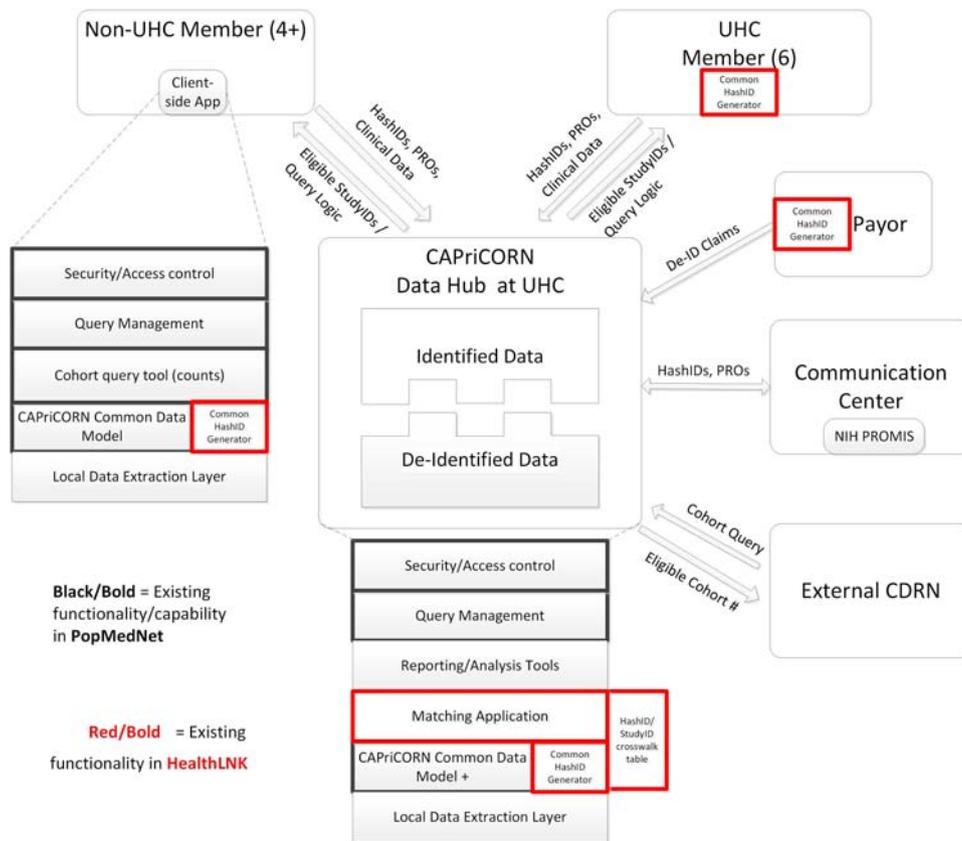
The majority of CAPriCORN sites exchange data with the University HealthSystem Consortium (UHC), an alliance of 120 academic medical centers and 301 of their affiliated hospitals representing the nation's leading academic medical centers (<http://www.uhc.edu>). Notably, five (LUHS, NM, RU, UI, UC) of our nine sites are current members of the UHC and exchange data for benchmarking purposes. Two additional sites (Alliance and CCHHS) are exploring plans to share data with UHC during the award period. The remaining sites (NS, JBVAMC and HinesVAH) will not share identifiable data with the UHC, and will manage their own data to support CAPriCORN activities (figure 1). All sites will work together to develop a common data model to facilitate efficient data sharing.

Building on the existing integration of identifiable data for five CAPriCORN institutions, UHC will work collaboratively with all CAPriCORN sites to facilitate development of the CAPriCORN Data Hub. This UHC role builds on work carried out by most of our institutions with Clinical and Translational Science Award funding through the Chicago Learning Effectiveness Research Network (Chicago LEARN) to develop and evaluate procedures to expand existing UHC utilization data to add laboratory values and other clinical data. The Data Hub will maintain segregation

of deidentified and identifiable data, with merging of data only for fully institutional review board (IRB)-approved protocols. It will also include data services such as security and access controls, query management, data reporting and analysis, patient matching, mapping to standard vocabularies, and provision of data to the nationally centralized repository for projects that are collaborative across the network. Although UHC already houses data from many of the CAPriCORN participants, a separate data mart (the CAPriCORN Data Hub) will be maintained for this project. The data will not be intermingled with other UHC (non-CAPriCORN) data. The data housed in the Data Hub will only be accessible to approved CAPriCORN participants. The non-UHC member institutions will serve as separate nodes within our network and will provide only deidentified data to the Data Hub.

Since 2009, informatics researchers at NM, UC, CCHHS, the Alliance, RU, and the UI have collaborated to develop a technical framework and a software application for deduplication and integration of deidentified, patient-level clinical data across institutions (HealthLNK). Our initial work merged clinical data on over 970 000 unique patients within the City of Chicago and over a total of 3.3 million patients within the Chicago metropolitan area.<sup>4</sup> With seed funding through the Otho S. A. Sprague Memorial Institute, a portion of the deidentified aggregated health data are shared through a public-facing website (<http://www.chicagohealthatlas.org>) hosted by CCT. As a proof-of-principle for our application, we included data from additional partners (LU) and updated clinical data from existing partners to identify a preliminary foundational cohort of over 5.7 million unique patients with an estimated 1.2 million unique primary care patients.

As part of the Data Hub, we will link two existing software tools, PopMedNet and the hashing/matching functions of HealthLNK (highlighted in black and red, respectively, in figure 1) to ensure rapid progress. PopMedNet is a software tool to enable distributed data queries, developed by researchers at Harvard Pilgrim with funding from the Agency for Healthcare Research and Quality.<sup>5</sup> The PopMedNet platform creates a secure framework for distributing data queries across multiple institutions and the flexibility to map data elements to a customized data model. Most instances of PopMedNet retrieve aggregated counts of eligible study cohorts without specifically returning unique identifiers that would enable disambiguation of patients across institutions.



**Figure 1** Overview of infrastructure and data flow. All participating sites will use a common data model. Distributed queries managed using PopMedNet, with distributed patient IDs (Hash-IDs) assigned using existing HealthLNK software. Deidentified data stored separately from identified data to accommodate different data use cases, and access managed by a central institutional review board (IRB).

The current version of the software in HealthLNK, refined recently by teams at three CAPriCORN institutions (NM, UI, UC), can generate a reliably disambiguated and consistent unique key, the Hash-ID, for patients based on commonly available identifiers (last name, first name, date of birth, gender, and social security number). This software creates a complementary function to PopMedNet—that is, a means to assign a common identity in a distributed fashion. This can be a significant informatics contribution to the wider PCORNet.

With these foundational tools in place, institutions will divide efforts into local versus centralized functions of the Data Hub by one of two methods:

1. Each institution contributing data to the Data Hub (UHC member institutions) will provide data security and access controls, query management services, cohort query tools,

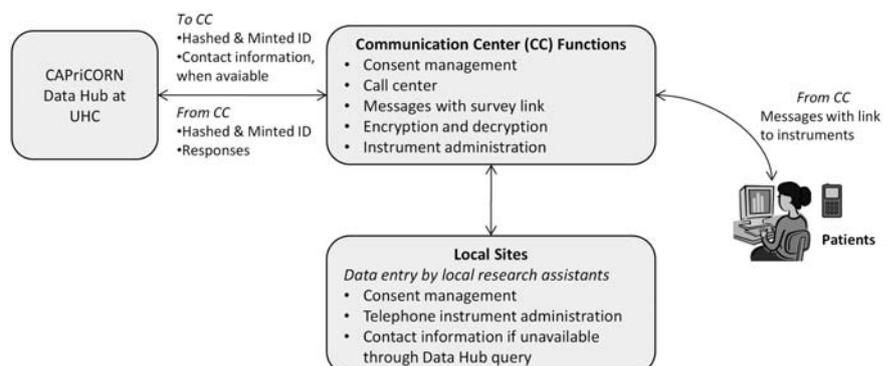
compliance with a CAPriCORN Common Data Model, Hash-ID services, and local data extraction tools.

2. The non-UHC member institutions will serve as separate nodes and provide deidentified data to the Data Hub. The non-UHC nodes will provide Hash-ID services, query management services, mapping of data to the CAPriCORN Common Data Model, and reporting and analysis tools for their own data.

### CENTRALIZED COMMUNICATION CENTER

The Data Hub will be integrated with our communications center (CC) to enable linking of clinical data with patient-reported and recruitment data. CAPriCORN will collect longitudinal data on patient-reported outcomes, symptom burden, functional status, behaviors, comorbidities, and detailed

**Figure 2** Process flow of the Communication Center. The Communication Center acts as a key interface between the CAPriCORN (Chicago Area Patient-Centered Outcomes Research Network) Data Hub and on-site primary research data collection.



epidemiological assessments. Rapid, prospective data collection will occur through a centralized CC linked to the Data Hub and through on-site primary data collection by research coordinators shared across multiple projects (figure 2). CAPriCORN sites will interact with the CC through a web-based platform, which will provide consent management, reporting tools, and a connection back to the Data Hub for construction of the final analytical dataset. Installation of the National Institute for Health's (NIH) PROMIS<sup>6</sup> (Patient-Reported Outcomes Measurement Information System) Assessment Center software system will enable administration of computer-adaptive testing instruments, which can be embedded into modules for administration of other standard instruments. The goal of the CC is to efficiently collect patient-generated data for clinical trials, observational studies, and comparative effectiveness evaluations. The CC will build web forms of selected instruments. For sophisticated computer-adaptive testing instruments, the CC will provide the infrastructure to collect data through computer-assisted telephone interviews and by sending short message service (SMS) texts and email messages to consenting patients. To minimize patient risk, identifiers will be encrypted during transmission and at rest while stored in the CC servers.

### SECURITY AND PRIVACY

CAPriCORN is committed to ensuring secure data environments that will protect individual privacy and to certifying that research using these data meets ethical standards for human subject protection. Stringent policies for security of computing environments, data, and compliance with applicable laws and regulations are in place at all institutions. These policies address administrative, technical, and physical security of computing environments. In addition, all institutions have policies, procedures, and practices designed to safeguard the use of clinical information for research purposes. As the CAPriCORN CDRN evolves, we will expand development of software and processes to support production of secure hash-based deidentified clinical datasets to allow patient disambiguation across disparate healthcare sites,<sup>7</sup> while ensuring data security for patient data in CAPriCORN.

CAPriCORN's human subject protection programs (HSPPs) will complement our data security practices. HSPPs are charged with ensuring that research meets ethical standards, is compliant with regulations, and is reviewed in a timely and collegial manner. The HSPP for CAPriCORN follows a consortium model, including: (1) a central IRB to review and oversee human subject research; (2) IRB representation from each institution with expertise encompassing the scope of proposed research activities, including the conduct of CER; (3) development of a unified electronic submission process and consistent consent procedures; (4) appointment of a dedicated expedited reviewer to direct the review of minimal risk protocols; and (5) establishment of an HSPP advisory board to monitor the quality of the review process and interface with the community advisory group and informatics group to discuss ethical and regulatory aspects of the design and implementation of the data repository.

### PARTICIPATORY RESEARCH: PATIENT AND CLINICIAN STAKEHOLDER ENGAGEMENT

Building on the principles of community-based participatory research<sup>8</sup> and models for stakeholder engagement in CER,<sup>9</sup> we established the Patient–Clinician Advisory Committee (PCAC). The PCAC provides the opportunity for patients and community-based clinicians to engage with the network to

provide guidance about (1) governance of data use, (2) research priorities, and (3) processes to incorporate patient and clinician perspectives regarding the design, implementation, and reporting of research and related results. The PCAC meets as a joint committee to guide research priorities and has voting representation on other key CAPriCORN committees, including the steering committee, human subjects working groups, and condition-specific working groups.

### SUMMARY

CAPriCORN is the Chicago-area CDRN focused on patient-centered outcomes research and designed to model characteristics required for development of the national PCORnet. CAPriCORN represents a population that is diverse in terms of ethnic and socioeconomic status with significant disparities in terms of access to and utilization of healthcare. Through its unique array of partners, including leading area medical institutions served by robust EHR systems and community partners engaged in health promotion and healthcare in the area, CAPriCORN will establish a network to share data, identify a cohort of over one million patients in the Chicago area, and establish a sustainable platform for patient-centered outcomes research that engages clinicians, patients, and other stakeholders in carrying out meaningful research.

### Author affiliations

- <sup>1</sup>Department of Medicine, Division of General Internal Medicine, Northwestern University, Chicago, Illinois, USA
- <sup>2</sup>Department of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA
- <sup>3</sup>Edward Hines Jr. VA Hospital, Chicago, Illinois, USA
- <sup>4</sup>Clinical Research Informatics, NorthShore University HealthSystem, Chicago, Illinois, USA
- <sup>5</sup>Health Sciences Division, Loyola University Health System, Chicago, Illinois, USA
- <sup>6</sup>Department of Medicine, Rush University Medical Center, Chicago, Illinois, USA
- <sup>7</sup>University of Illinois Hospital & Health Sciences System, Chicago, Illinois, USA
- <sup>8</sup>Department of Medicine, Cook County Health and Hospital System, Chicago, Illinois, USA
- <sup>9</sup>VA Information Resource Center, Edward Hines Jr VA Hospital, Chicago, Illinois, USA
- <sup>10</sup>Alliance of Chicago Community Health Services, Chicago, Illinois, USA
- <sup>11</sup>Department of Medicine, Section of Hospital Medicine, University of Chicago, Chicago, Illinois, USA
- <sup>12</sup>Department of Pediatrics, University of Chicago, Chicago, Illinois, USA
- <sup>13</sup>Clinical Directors Network (CDN) and The Rockefeller University Center for Clinical and Translational Science, New York City, New York, USA
- <sup>14</sup>University HealthSystem Consortium, Chicago, Illinois, USA
- <sup>15</sup>Tufts Medical Center, Boston, Massachusetts, USA
- <sup>16</sup>Illinois Medical District Commission, Chicago, Illinois, USA
- <sup>17</sup>Chicago Community Trust, Chicago, Illinois, USA

**Acknowledgements** We gratefully acknowledge Jörn Boehnke, John Eric Humphries, and Scott Kominers for their work on optimization of the hashing and matching software.

**Contributors** ANK created the initial regional data aggregation, led the design of the informatics infrastructure and workgroup, facilitated development of the collaborative, and drafted and revised the paper. DMH co-led the informatics workgroup, contributed to the design of the informatics infrastructure, and drafted and revised the paper. SG contributed to the initial regional data aggregation, collected and collated data across institutions, and drafted and revised the paper. AES led data collection and organization across multiple institutions, contributed to the design of the informatics infrastructure, and edited and revised the paper. RP participated in the initial regional data aggregation, contributed to the design of the informatics infrastructure, and revised the paper. BH participated in the initial regional data aggregation, contributed to the design of the informatics infrastructure, and revised the paper. SAS participated in the initial regional data aggregation, contributed to the design of the informatics infrastructure, and revised the paper. NB contributed to the design of the informatics infrastructure and revised the paper. FA contributed to the design of the communication center and revised the paper. WET led the design of the communication center, participated in the initial regional data aggregation, and revised the paper. ET contributed to the design of the informatics infrastructure and revised the paper. FDR participated in the initial regional data aggregation, facilitated development of

the collaborative, and revised the paper. AH contributed to the design of the informatics infrastructure and revised the paper. EOK participated in the initial regional data aggregation and revised the paper. SB contributed to the design of the informatics infrastructure and revised the paper. SLV contributed to the design of the informatics infrastructure and revised the paper. JCS facilitated development of the collaborative, contributed to the design of the informatics infrastructure, and revised the paper. JNT facilitated development of the collaborative and revised the paper. MAS contributed to the design of the informatics infrastructure and revised the paper. DL contributed to the design of the informatics infrastructure and revised the paper. JBW facilitated development of the collaborative and revised the paper. RHK facilitated development of the collaborative, led creation of the overall budget, and revised the paper. JAK facilitated development of the collaborative and revised the paper. DOM participated in the initial regional data aggregation, facilitated development of the collaborative, and revised the paper. JMC led overall administrative coordination of the collaborative, facilitated development of the collaborative, and revised the paper. TM facilitated development of the collaborative and revised the paper.

**Funding** This work was supported by a grant from the Otho S. A. Sprague Memorial Institute. DMH is supported by a Department of Veterans Affairs (VA) Research Career Scientist Award RCS-98-352. Early testing of this network was supported in part by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grants UL1TR000050 and UL1TR000430. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the VA. Chicago LEARN was supported through a Comparative Effectiveness Research Supplement to 'Translational Research at the University of Chicago' NIH UL1 (Solway, PI, Meltzer Project Leader), and the Patient-Centered Outcomes Research Institute (Patient Navigator to Reduce Readmissions, IH-12-11-4365, Krishnan, PI).

**Competing interests** None.

**Provenance and peer review** Commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

## REFERENCES

- 1 National Center for Health Statistics. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). <http://www.cdc.gov/nchs/icd/icd9cm.htm> (accessed 20 Sep 2013).
- 2 American Medical Association. Current Procedural Terminology. <http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.page?> (accessed 20 Sep 2013).
- 3 Forrey AW, McDonald CJ, DeMoor G, *et al.* Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clin Chem* 1996;42:81–90.
- 4 Kho AN, Cashy JP, Hota BN, *et al.* The Chicago Health Atlas: A public resource to visualize health conditions and resources in Chicago. AMIA Theater Style Demonstration. AMIA Symposium; 2012.
- 5 Brown JS, Holmes JH, Shah K, *et al.* Distributed Health Data Networks: A practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. *Med Care* 2010;48:S45–S5.
- 6 Gershon RC, Rothrock N, Hanrahan R, *et al.* The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *J Appl Meas* 2010;11:304.
- 7 Galanter WL, Applebaum A, Boddipalli V, *et al.* Migration of patients between five urban teaching hospitals in Chicago. *J Med Syst* 2013;37:9930.
- 8 Shalowitz M, *et al.* Community-based participatory research: a review of the literature with strategies for community engagement. *J Dev Behav Pediatr* 2009;30:350–61.
- 9 Deverka P, Lavalley D, Tunis S, *et al.* Stakeholder participation in comparative effectiveness research: defining a framework for effective engagement. *J Comp Eff Res* 2012;1:181–94.