PEDSnet: a National Pediatric Learning Health System

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1136/amiajnl-2014-002743

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:12717617

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
PEDSnet: a National Pediatric Learning Health System


ABSTRACT

A learning health system (LHS) integrates research done in routine care settings, structured data capture during every encounter, and quality improvement processes to rapidly implement advances in new knowledge, all with active and meaningful patient participation. While disease-specific pediatric LHSs have shown tremendous impact on improved clinical outcomes, a national digital architecture to rapidly implement LHSs across multiple pediatric conditions does not exist. PEDSnet is a clinical data research network that provides the infrastructure to support a national pediatric LHS. A consortium consisting of PEDSnet, which includes eight academic medical centers, two existing disease-specific pediatric networks, and two national data partners form the initial partners in the National Pediatric Learning Health System (NPLHS). PEDSnet is implementing a flexible dual data architecture that incorporates two widely used data models and national terminology standards to support multi-institutional data integration, cohort discovery, and advanced analytics that enable rapid learning.

BACKGROUND

Most pediatric chronic conditions are rare diseases, requiring multi-institutional research networks to conduct research and learn how variations in clinical practice affect pediatric health outcomes.1–3 A learning health system (LHS) integrates clinical studies done in routine care settings, leverages structured data capture at every encounter, and incorporates quality improvement methods to implement advances in new knowledge and care delivery, with active and meaningful patient participation.4–8 While disease-specific examples of pediatric LHSs have shown tremendous improvement in clinical outcomes,9–11 a national digital architecture to support the rapid implementation of LHSs across multiple pediatric conditions does not exist.12,13

PEDSnet is a clinical data research network (CDRN) that forms the digital infrastructure to support a national pediatric LHS. The PEDSnet CDRN is a consortium of eight children’s hospitals, two existing patient-centered disease-specific pediatric networks addressing inflammatory bowel disease and complex congenital heart disease, a newly formed pediatric obesity network, and two national data partners. Together they form the initial components of the National Pediatric Learning Health System (NPLHS). The NPLHS will establish the data sharing environment to enable a community of patients and clinicians, interacting at the point of care, to generate data that can be repurposed for research and quality improvement and to support continuous monitoring of outcomes that identify specific management practices as targets for comparative effectiveness research (CER). This brief report describes the PEDSnet CDRN digital infrastructure.

The PEDSnet CDRN is one of 11 CDRNs funded by the Patient Centered Outcomes Research Institute (PCORI) as part of the newly established PCORnet, a national patient-centered clinical research network for conducting clinical effectiveness research (see http://www.pcornet.org).14 PEDSnet consists of eight academic pediatric health centers: Children’s Hospital of Philadelphia (CHOP), Cincinnati Children’s Hospital Medical Center (CCHMC), Children’s Hospital Colorado, Nemours Children’s Health System, Nationwide Children’s Hospital, St. Louis Children’s Hospital, Seattle Children’s Hospital, and Boston Children’s Hospital. With regional catchment areas extending across 22 states, and a national and international specialty referral base, PEDSnet hospitals provide care for 2.1 million patients, or 2.8% of all children in the USA, from diverse demographic, geographic, and socioeconomic backgrounds (table 1). Additional participants include the ImproveCareNow (ICN) Inflammatory Bowel Disease Network,9 the National Pediatric Cardiology Quality Improvement Collaborative (NPCQIC) which focuses on complex congenital heart disease,15 and a newly formed Healthy Weight obesity prevention network. Partnerships with Express Scripts, a national pharmacy benefits management company, and IMS Health, a data aggregator of multi-payer administrative databases, extends the reach of NPLHS beyond the core PEDSnet CDRN institutions.

THE PEDSnet CDRN DIGITAL INFRASTRUCTURE

PEDSnet institutions use a variety of clinical systems from multiple vendors, including internal clinical registries (table 2). Differences in implementation and local documentation practices prevent institutions from sharing data, even among institutions that have installed the same vendor product.

Interoperable dual common data models and environments

The informatics teams in PEDSnet institutions have substantial expertise in two open-source data-sharing platforms—c2b2, developed by Harvard University,16 and OMOP, created by the Observational Medical Outcomes Partnership.17
To leverage existing institutional investments and capabilities in i2b2 and OMOP, PEDSnet will support data submissions using either i2b2 (figure 1, top pathway) or OMOP (figure 1, bottom pathway). The PEDSnet DCC will develop conversion routines for replicating i2b2 data into OMOP and vice versa. Institutions can provide data in whichever format fits local data efforts. The burden of transforming data into the other data model will be managed centrally at the PEDSnet-CDRN DCC, which is responsible for data management, quality assessment, and network-based data analytics. Consequently, each site will have access to two versions of its data (i2b2 and OMOP format), both of which will conform to national terminology standards.

**Terminology harmonization**

Differences in data structures and codes across institutional electronic health records (EHRs) are a significant barrier to multi-institutional data sharing.\(^{18-21}\) PEDSnet will require institutions to harmonize variables and value sets using terminologies proposed by the US government and health care community (eg, current Department of Health and Human Services (DHHS) Meaningful Use certification criteria, such as SNOMED-CT, ICD, LOINC, RxNorm) and other terminology standards that support data exchange (eg, CDISC ODM and the HL-7 consolidated CDA).\(^{22-24}\) The National Library of Medicine (NLM), under contract with DHHS, supports the Value Set Authority Center (VSAC), which serves as the national resource for defining all terms that will be required for quality reporting from EHRs.\(^{25}\) Similarly, the NLM Common Data Element Resource Portal provides a central resource for NIH-sponsored initiatives to support interoperable clinical research.\(^{26}\)

**Linking to the NIH pediatric terminology effort**

Pediatric concepts are poorly represented in existing terminology standards. The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) supports a pediatrics terminology harmonization effort focused on improving the coverage of critical pediatric research terms in clinical terminologies such as SNOMED-CT.\(^{27}\) Patient-centered registries often develop highly disease-specific questions, responses, and outcomes. For example, the ICN registry contains disease-specific questions and responses, such as ‘periarectal disease at current exam’ or ‘mildly inflamed skin tag(s), 1–2 indolent fistulas with scant drainage (no tenderness)’, that are required for adequate description of specific conditions. These terms, however, are unlikely to be found in any standard terminology and therefore are more difficult to harmonize with relevant data available from partners outside of ICN. PEDSnet will partner with the NIH Pediatric Terminology effort to submit new

<table>
<thead>
<tr>
<th>Total number of patients</th>
<th>2,171,000 (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>0–2 years (up to 23 months)</td>
<td>391,000 (18%)</td>
</tr>
<tr>
<td>2–5 years (24–59 months)</td>
<td>412,000 (19%)</td>
</tr>
<tr>
<td>5–11 years (60–43 months)</td>
<td>673,000 (31%)</td>
</tr>
<tr>
<td>12–17 years (144–215 months)</td>
<td>499,000 (23%)</td>
</tr>
<tr>
<td>18–24 years (216–287 months)</td>
<td>109,000 (5%)</td>
</tr>
<tr>
<td>25+ years (288 months+)</td>
<td>87,000 (4%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>No. patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>1,042,000 (48%)</td>
</tr>
<tr>
<td>Male</td>
<td>1,129,000 (52%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race/ethnicity</th>
<th>No. patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>43,000 (2%)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>391,000 (18%)</td>
</tr>
<tr>
<td>Hispanic or Latino ethnicity</td>
<td>174,000 (8%)</td>
</tr>
<tr>
<td>White</td>
<td>1,216,000 (56%)</td>
</tr>
<tr>
<td>Other/unknown</td>
<td>499,000 (23%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic health center</th>
<th>No. patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston Children’s Hospital</td>
<td>235,000 (11%)</td>
</tr>
<tr>
<td>Children’s Hospital of Philadelphia</td>
<td>435,000 (20%)</td>
</tr>
<tr>
<td>Cincinnati Children’s Hospital Medical Center</td>
<td>231,000 (11%)</td>
</tr>
<tr>
<td>Children’s Hospital Colorado</td>
<td>185,000 (9%)</td>
</tr>
<tr>
<td>Nationwide Children’s Hospital</td>
<td>304,000 (14%)</td>
</tr>
<tr>
<td>Nemours Children’s Health System</td>
<td>248,000 (11%)</td>
</tr>
<tr>
<td>Seattle Children’s Hospital</td>
<td>153,000 (7%)</td>
</tr>
<tr>
<td>St. Louis Children’s Hospital</td>
<td>379,000 (17%)</td>
</tr>
</tbody>
</table>

Data were extracted from electronic health records during calendar year 2012.

The i2b2 data model and integrated cohort discovery application has been designed to be ‘data-browser friendly’, allowing individuals to explore high level concepts and identify cohorts of interest, but it does not have a library of highly detailed analyses. The OMOP data model and analytic methods applications are more ‘analysis-focused’, with a library of highly complex queries involving numerous variables and relationships between variables. Thus, the two data models, applications, and tools are complementary with i2b2 supporting cohort discovery and patient recruitment and OMOP supporting data analytics for CER. PEDSnet is implementing a centralized data-coordinating center (DCC) for aggregating specified data elements across institutions. PEDSnet may also support distributed queries when that strategy is more suitable for particular research questions.

**Linking to the NIH pediatric terminology effort**

Pediatric concepts are poorly represented in existing terminology standards. The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) supports a pediatrics terminology harmonization effort focused on improving the coverage of critical pediatric research terms in clinical terminologies such as SNOMED-CT. Patient-centered registries often develop highly disease-specific questions, responses, and outcomes. For example, the ICN registry contains disease-specific questions and responses, such as ‘periarectal disease at current exam’ or ‘mildly inflamed skin tag(s), 1–2 indolent fistulas with scant drainage (no tenderness)’, that are required for adequate description of specific conditions. These terms, however, are unlikely to be found in any standard terminology and therefore are more difficult to harmonize with relevant data available from partners outside of ICN. PEDSnet will partner with the NIH Pediatric Terminology effort to submit new
pediatric terms and their definitions to national and international standards bodies, such as the NLM and the International Health Terminology Standards Development Organization (IHTSDO; http://www.ihtsdo.org), the organization responsible for the development of SNOMED-CT.

The PEDSnet common data models will have breadth (across all 2 million children in NPLHS) and depth (detailed clinical data, patient-reported outcome (PRO) data, and administrative data for the three related disease cohorts). The EHR-derived data will include structured patient demographics, anthropometrics, physiological measurements, diagnoses (mapped to SNOMED-CT), medications (mapped to RxNorm), laboratory results (mapped to LOINC), procedures, and specialty-specific observations (mapped to standard terminologies when terms already exist). PEDSNet will implement an opportunistic terminology mapping approach, starting with most widely used laboratory and physiological measurement concepts across the sites and all clinical concepts relevant to research use cases. Future CER studies may require additional data elements, which will be mapped at that time.

PATIENT REPORTED OUTCOMES

The digitization of healthcare is transforming the potential for clinical research as millions of patients and encounters become available through electronic records. It is less clear, however, how patients’ preferences and perspectives on health, well being, and behavior will be obtained or utilized. Until recently, virtually all PROs were recorded using paper. Today, many electronic data capture systems, including interactive voice response, handhelds, tablets, and other web-enabled technologies, exist but are not used widely in clinical settings. While progress on electronic data capture systems, including interactive voice response, handhelds, tablets, and other web-enabled technologies, exists, much research work will include extending patient interactions into alternative media, such as web applications and SMS text messaging, which may be more effective methods for obtaining PROs from a high proportion of participants.

NATIONAL DATA PARTNERS

PEDSNet has established new relationships with two national companies that manage pharmacy benefits (Express Scripts) and multi-payer administrative databases (IMS Health). Express Scripts is a Fortune 100 company that processes pharmacy claims for 100 million Americans. Data from Express Scripts provides information on medicines actually dispensed to the patient, whereas EHR data provides information on medications prescribed. Knowing which medications are prescribed and dispensed provides valuable information about patients’ adherence to medical regimens. A pilot study on inflammatory bowel disease patients seen at St. Louis Children’s Hospital found that Express Scripts data could be linked to EHR data for 60% of those patients. IMS Health is one of the world’s largest providers of healthcare ‘big data’. It integrates data from a variety of sources, such as medical, hospital, and other healthcare records sourced from multiple data suppliers. These databases contain

patients to complete PRO measures in waiting rooms (on tablets or kiosks) or on their home computers via a patient portal. Surveys can be tied to events, such as an upcoming visit or a scheduled surgery, simplifying the physician’s use of ePROs in follow-up (eg, a physician could ‘order’ a pain PRO 1 week pre-surgery, 1 week post, and 3 months post). Five of the eight PEDSNet-CDRN institutions will have the Epic PRO functionality available by the end of CY2013. The NIH PROMIS (http://www.nihpromis.org) measures for adults and children can be accessed using this technology.

Research Electronic Data Capture system (REDCap): REDCap, from Vanderbilt University, can be used to collect self-reported ePROs and other electronic patient-reported data via questionnaires. Boston Children’s Hospital has created an interface that converts data from REDCap into i2b2 format that can be used to integrate patient-reported data obtained using REDCap into the PEDSnet CDRN. These two systems will enable all PEDSNet-CDRN institutions to incorporate ePROs during in-clinic visits or through online recruitment outside the clinic. In preliminary work, we have found that the best approach for collecting high rates of PRO data is to obtain them in the waiting areas before clinic visits, rather than via a pre-visit email. PEDSNet’s patient engagement work will include extending patient interactions into alternative media, such as web applications and SMS text messaging, which may be more effective methods for obtaining PROs from a high proportion of participants.

Figure 1 Overall architecture and data flows in PEDSnet, highlighting dual methods (i2b2, OMOP) for data submission and sharing.

over 16 billion records of standardized, patient-level, de-identified healthcare encounter records representing all 50 states and more than 260 million patients worldwide. Data are linked at the patient-level across data sources. One of the great strengths of the IMS Health database is that it is multi-payer and national in scope—that is, includes many health insurers within and across markets. Adding insurance claims to EHR data will provide insights on use of services outside the children’s hospitals, including use of outside specialists and emergency departments.

CONCLUSION
PEDSnet will transform pediatric healthcare and children’s health by developing an extensible and efficient digital infrastructure that enables all participants to collaborate in the work of producing new knowledge and improving health and care delivery. PEDSnet benefits from robust pre-existing resources and a unique history of collaboration by children’s hospitals that has fundamentally reshaped outcomes for previously fatal diseases, such as cystic fibrosis and many childhood cancers. By developing a dual common data model platform with 2h2 and OMOP, PEDSnet allows organizations to participate in the data-sharing network based on existing institutional priorities and programs. By contributing to pediatric terminology standards activities, PEDSnet will help expand the availability and interoperability of complex pediatric concepts that are not present in existing terminologies. As the underlying digital infrastructure to a LHS, PEDSnet will enable the rapid implementation of new evidence into clinical practice and will address fundamental questions of clinical effectiveness for children and their families, particularly for individuals affected by serious, and generally rare, illness that persists into adulthood.

Author affiliations
1 Department of Pediatrics, Children’s Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA
2 Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
3 Department of Pediatrics, The Anderson Center for Health Systems Excellence, Cincinnati Children’s Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA
4 Department of Pediatrics, Division of Biomedical Informatics, Cincinnati Children’s Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA
5 Seattle Children’s Hospital, Seattle, Washington, USA
6 Department of Pediatrics, University of Washington School of Medicine, Seattle, Washington, USA
7 Division of General Pediatrics, Boston Children’s Hospital, Boston, Massachusetts, USA
8 Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA
9 Nemours Children’s Hospital and the College of Medicine, University of Central Florida, Orlando, Florida, USA
10 Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children’s Hospital and Departments of Pediatrics and Statistics, The Ohio State University, Columbus, Ohio, USA
11 Department of Pathology and Laboratory Medicine, Children’s Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA
12 Department of Pediatrics, St. Louis Children’s Hospital, Washington University School of Medicine, St. Louis, Missouri, USA
13 Department of Pediatrics, Children’s Hospital Colorado, University of Colorado and the Colorado Clinical and Translational Sciences Institute, Aurora, Colorado, USA

Contributors All authors meet all four ICMJE criteria for authorship. MGK was responsible for the initial draft manuscript that was reviewed, edited, and expanded by all listed authors. MGK was responsible for the final draft that was reviewed and approved by all authors prior to submission. Descriptive and quantitative data in tables 1 and 2 were provided by the authors from those institutions. They are responsible for the accuracy of these data.

Funding This work was supported by PCORI Contract CDRN-1306-01556 (All), AHRQ 1R01HS022974 Enhancing the Sustainability of a Pediatric Learning Health System (PAM); AHRQ 1R01HS019912 Scalable Partnering Network for CER: Across Lifespan, Conditions, and Settings (MGK) and NIH/NCATS Colorado CTSG Grant Number UL1 TR001082 (MGK).

Competing interests None.

Provenance and peer review Commissioned; internally 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
6 Slutsky JR. Moving closer to a rapid-learning health care system. Health Aff (Millwood) 2007;26:w122–4.
11 Margolis PA, Peterson LE, Seid M. Collaborative Chronic Care Networks (C3Ns) to transform chronic illness care. Pediatrics 2013;131(suppl 4):S219–23.
Brief communication


