



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Query Health: standards-based, cross-platform population health surveillance

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Klann, Jeffrey G, Michael D Buck, Jeffrey Brown, Marc Hadley, Richard Elmore, Griffin M Weber, and Shawn N Murphy. 2014. "Query Health: standards-based, cross-platform population health surveillance." Journal of the American Medical Informatics Association : JAMIA 21 (4): 650-656. doi:10.1136/amiajnl-2014-002707. http://dx.doi.org/10.1136/amiajnl-2014-002707 .
Published Version	doi:10.1136/amiajnl-2014-002707
Accessed	February 16, 2015 4:53:16 PM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:12717615
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

(Article begins on next page)



OPEN ACCESS

Query Health: standards-based, cross-platform population health surveillance

Jeffrey G Klann,^{1,2,3} Michael D Buck,⁴ Jeffrey Brown,^{2,5} Marc Hadley,⁶ Richard Elmore,⁷ Griffin M Weber,^{2,8} Shawn N Murphy^{1,2,3}

For numbered affiliations see end of article.

Correspondence to

Dr Jeffrey G Klann, Research Computing, Partners Healthcare System, One Constitution Center, Charlestown, MA 02129, USA; jeff.klann@mgh.harvard.edu

Received 5 February 2014

Accepted 11 March 2014

Published Online First

3 April 2014

ABSTRACT

Objective Understanding population-level health trends is essential to effectively monitor and improve public health. The Office of the National Coordinator for Health Information Technology (ONC) Query Health initiative is a collaboration to develop a national architecture for distributed, population-level health queries across diverse clinical systems with disparate data models. Here we review Query Health activities, including a standards-based methodology, an open-source reference implementation, and three pilot projects.

Materials and methods Query Health defined a standards-based approach for distributed population health queries, using an ontology based on the Quality Data Model and Consolidated Clinical Document Architecture, Health Quality Measures Format (HQMF) as the query language, the Query Envelope as the secure transport layer, and the Quality Reporting Document Architecture as the result language.

Results We implemented this approach using Informatics for Integrating Biology and the Bedside (i2b2) and hQuery for data analytics and PopMedNet for access control, secure query distribution, and response. We deployed the reference implementation at three pilot sites: two public health departments (New York City and Massachusetts) and one pilot designed to support Food and Drug Administration post-market safety surveillance activities. The pilots were successful, although improved cross-platform data normalization is needed.

Discussions This initiative resulted in a standards-based methodology for population health queries, a reference implementation, and revision of the HQMF standard. It also informed future directions regarding interoperability and data access for ONC's Data Access Framework initiative.

Conclusions Query Health was a test of the learning health system that supplied a functional methodology and reference implementation for distributed population health queries that has been validated at three sites.

BACKGROUND AND SIGNIFICANCE

The Institute of Medicine has developed a vision for a learning health system (LHS), which will integrate the nation's electronic healthcare components to share and learn from each other.¹ An initial step toward this is the Meaningful Use Incentive Program (MU), which encourages the adoption and use of electronic health systems. This is laying groundwork for LHS, which will cross organizational boundaries for tasks such as comparative effectiveness research, population health surveillance, and dissemination of evidence-based medicine. As the Institute of Medicine describes it: 'The increased complexity of health care requires a sustainable system that gets the right care to the right

people when they need it, and then captures the results for improvement. The nation needs a healthcare system that learns.'²

The Office of the National Coordinator for Health Information Technology (ONC) has embraced LHS in their strategic plan.³ As a large-scale test of LHS functionality, they launched the Query Health initiative in September 2011. Query Health is a public-private collaboration to develop standards and services to enable distributed, secure, standards-based population health measurement.⁴ This capability to measure population-level health trends is essential to public health.

The ONC embraced a distributed query model in Query Health, which eliminates centralization of data by 'bringing questions to the data.' Individual organizations process queries and disclose only the minimum necessary information to answer the query—often aggregate statistics—avoiding many privacy and security concerns. This federated approach—which requires deep content and system knowledge of the contributing health systems⁵—is nonetheless being used effectively for research, cohort selection, and population health surveillance.^{6–14}

Over the past 2 years, Query Health has developed a methodology and a flexible, open-source reference implementation.^{15–24} We have piloted the implementation at several locations and collected feedback, which is guiding future work on a national scale.^{25–26}

OBJECTIVE

Query Health was tasked with three major goals. First, to define a methodology for distributed, secure, standards-based clinical queries. Existing standards were used wherever possible. Second, to develop a reference implementation using best-of-breed technological components.²⁷ Third, to implement components of this reference implementation at three pilot sites, to gauge the effectiveness of Query Health in real-world healthcare scenarios.²² Two of the pilots were in cooperation with Departments of Health (New York City and Massachusetts) for disease monitoring and surveillance. The third pilot focused on the potential to expand the data resources available for medical product safety surveillance through the Food and Drug Administration (FDA) Mini-Sentinel project.²⁸

MATERIALS AND METHODS

The general Query Health workflow is as follows: (1) investigators develop 'questions' to ask the data using a standard ontology and query format; (2) the question is securely distributed through a 'query envelope' to participating data partners; (3)



Open Access
Scan to access more
free content



CrossMark

To cite: Klann JG, Buck MD, Brown J, et al. *J Am Med Inform Assoc* 2014;**21**:650–656.

data partners process the query and return aggregate results; (4) results are combined and reported back to the investigator (figure 1).

Ontology

The Query Health ontology uses the Consolidated Clinical Document Architecture (C-CDA) to instantiate a hierarchical terminology based on the National Quality Forum's Quality Data Model (QDM).¹⁵ These standards are already required as part of the 2014 certification for stage 2 of MU, so implementers are encouraged to use these standards instead of deriving new models. C-CDA is used to produce reports of patient data, and QDM defines data elements required for clinical quality measures (CQMs).^{29 30} The reason for using this approach is summarized well in the 2014 Electronic Health Record (EHR) Certification Final Rule: 'this standard provides, for the first time, a method of moving a 'snapshot' of patient data from one EHR technology to another without loss of semantic integrity.'³¹

Query format

The Health Level 7 (HL7) Health Quality Measures Format (HQMF) is Query Health's 'question language'. HQMF is an XML standard for platform-neutral clinical queries, which already has national focus and adoption because of MU. The National Quality Forum and the Centers for Medicare and Medicaid Services are using it and have released many of their CQMs in HQMF format.^{32 33} It is possible that HQMF will appear in the MU stage 3 requirements.

Despite the national attention, HQMF had inadequate computability in 2011. Query Health worked with HL7 to develop a second revision of HQMF that balances the flexibility needed by query developers and the computational tractability needed in implementations. This revision will be available through HL7 shortly and will be used for future HQMF-based CQMs.³⁴

In addition to the new HQMF format, a forthcoming HQMF QDM implementation guide will enumerate implementation details of using the Query Health ontology with HQMF.³⁵

Query transport

The Query Health Query Envelope standard supports secure transport of queries and results through a distributed network. It provides very granular control to data partners to authorize or decline data release, and it is independent of query and result formats. It is a flexible, secure transport mechanism.³⁶

Results format

The HL7 Quality Reporting Document Architecture (QRDA) is Query Health's 'result language'. QRDA is another platform-neutral XML language based on the HL7 Reference Implementation Model and is already required for quality reporting in MU stage 2.³⁰

Evaluation

We implemented this methodology as a flexible reference implementation using adaptable, best-of-breed, open-source technologies. Figure 2 is a summary of this implementation. We then piloted the reference implementation at three sites using different components for each use case. We collected feedback on the implementation and pilot experiences. Table 1 is a summary of this evaluation.

RESULTS

Implementation

PopMedNet

For query transport, PopMedNet was selected. PopMedNet is an open-source distributed data-sharing platform funded by the Agency for Healthcare Research and Quality, the FDA, ONC, and the National Institutes of Health (NIH).^{16 37} It is a key component of several large-scale distributed networks, including the FDA Mini-Sentinel, the NIH Health Care System Research Collaboratory Distributed Research Network, and the Massachusetts Department of Public Health MDPHnet system.⁹⁻¹¹ PopMedNet will be used by the newly funded Patient-Centered Outcomes Research Institute (PCORI) National Patient-Centered Clinical Research Network (PCORnet) to help create and operate a 'network-of-networks' to support clinical research.³⁸ PopMedNet

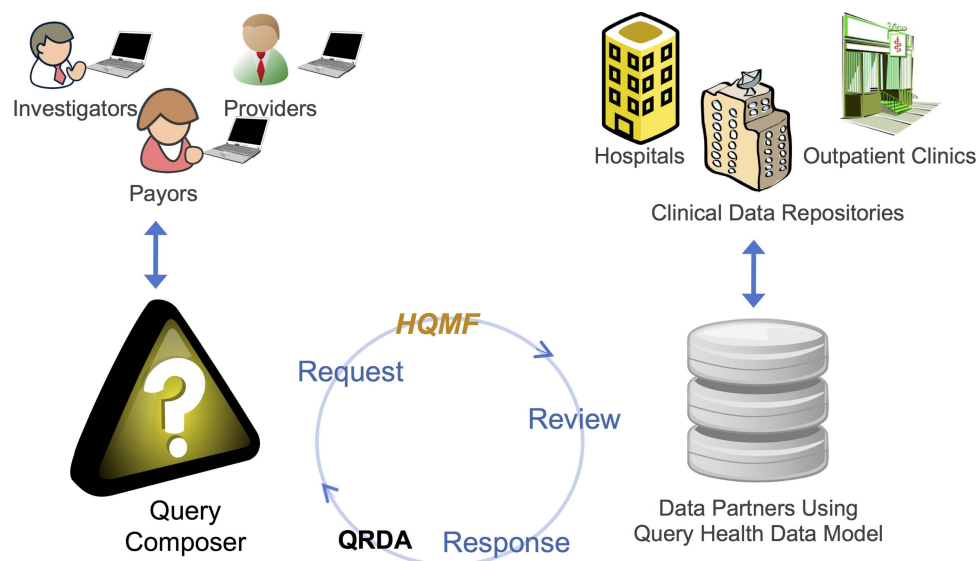
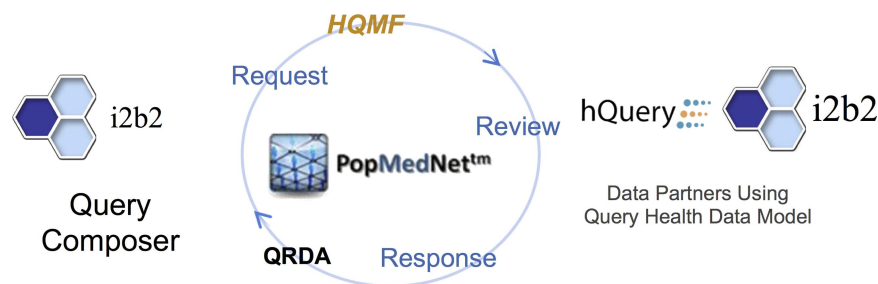


Figure 1 Overall design of Query Health. Various stakeholders can develop queries, which are distributed securely and sent to a variety of data partners. These data partners process the queries and return aggregate counts, so that sensitive data never leave individual sites. Query Health uses a variety of standards: a Query Envelope, a Data Model, Health Quality Measures Format (HQMF), and Quality Reporting Document Architecture (QRDA).

Figure 2 The technologies used in the Query Health reference implementation. Individual pilots varied in their use of these components (see Table 1). Informatics for Integrating Biology and the Bedside (i2b2) is used as a graphical query composer. PopMedNet is the query distribution and authentication engine. i2b2 and hQuery are the back-end data warehouses used by data partners to connect to Query Health. A Health Quality Measures Format (HQMF) is used to communicate queries in a standardized format.



defines a network topology, manages access controls, distributes queries to participating partner sites for local execution, and manages the query response. It is agnostic to query, response, and data formats, making it ideal for integration with disparate systems. The PopMedNet architecture enables synchronous or asynchronous distributed querying, enables partners to opt-in to each query or type of query (via an access control layer), and allows partners to use to their internal workflow for query response.

Each PopMedNet query includes query metadata that we extended to conform to the Query Health Query Envelope standards. Each query is defined at a central PopMedNet portal, securely distributed using the Query Envelope, ‘unpacked’ at the participating site, and processed locally by a Data Mart Client program that manages the local execution and response. The query response is also packaged into the Query Envelope for return to the requester via PopMedNet.

i2b2

For query composition and processing, Informatics for Integrating Biology and the Bedside (i2b2) was selected. i2b2 is an open-source clinical data analytics platform funded by the NIH and used at over 100 sites nationwide.^{39 40} It provides an intuitive, graphical, web-based query builder as well as a flexible analytical database design. Its component-based architecture makes it easily adaptable to new-use cases, and it has already been used in another distributed query platform called SHRINE, the Shared Health Research Information Network.⁶ Over a third of PCOR’s

PCORnet is currently using i2b2, including both Clinical Data Research Networks and Patient-Powered Research Networks. Harvard’s PCORnet Clinical Data Research Network involves the interoperability of i2b2 at 10 health systems.^{38 41}

i2b2 is a set of web service components, known as cells, that collectively make up a ‘hive’. It is possible to add and remove components for different use cases of i2b2. To Query Health-enable i2b2, we created three new cells:

- ▶ A PopMedNet client adapter, which sends investigator-developed queries to the PopMedNet web portal.
- ▶ A PopMedNet server adapter, which receives queries from and sends results back to PopMedNet.
- ▶ An HQMF translator, able to process HQMF revision 2.

These cells integrate into the hive. This is described below and shown in figures 3 and 4. The i2b2 platform and these cells are open-source.^{15 17}

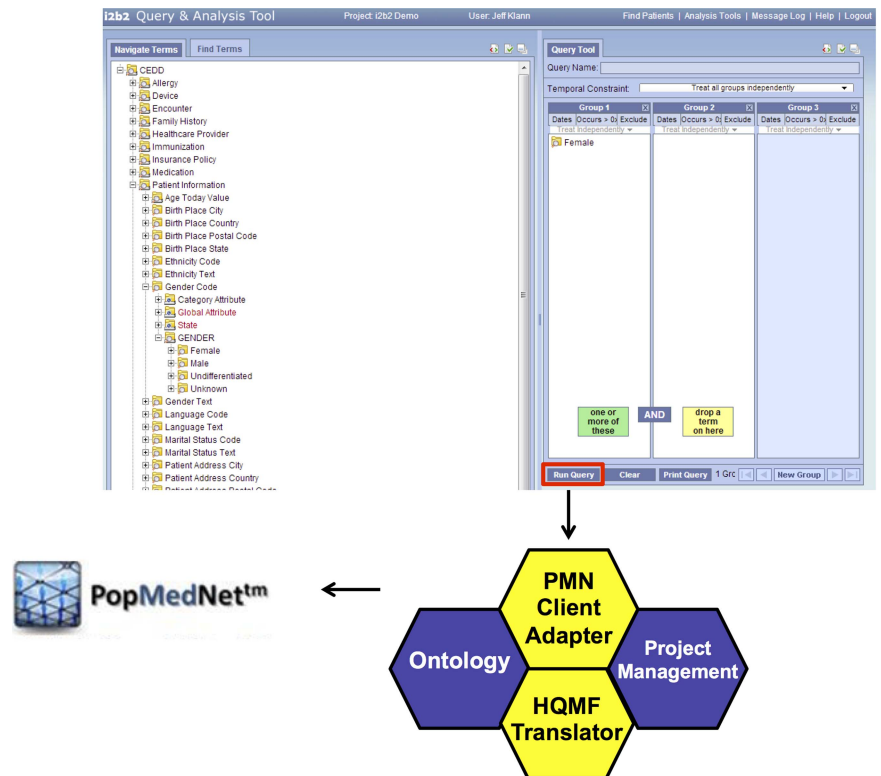
When an investigator develops a query in the Query Health i2b2 query builder, the query is sent to the PopMedNet Client Adapter rather than the local data repository. This Adapter then transmits the query (which is in an i2b2-compliant XML format) to the HQMF translator, and it sends the resulting HQMF query onward to the PopMedNet portal. The query builder then displays results as they arrive in the ‘previous queries’ window (figure 3).

When a Query Health-enabled i2b2 system is sent an HQMF query by the PopMedNet Data Mart Client, it is routed to the PopMedNet Server Adapter, which similarly sends the query to

Table 1 Summary of the three Query Health pilots

Pilot site	Goal and feedback	Technology	Status
New York City Department of Health and Mental Hygiene	<i>Initial goal:</i> Demonstrate standards-based vendor-neutral distributed analytics solution using Query Health Standards <i>Current goal:</i> Launch health information exchange-based solution to obtain aggregate city-wide healthcare statistics <i>Feedback:</i> Standards-based aggregate distributed analytic solutions are now capable of delivering valuable results. Challenges remain in cross-site data harmonization	<i>Composer:</i> i2b2 <i>Envelope:</i> PopMedNet <i>Processor:</i> i2b2 <i>Other:</i> HQMF, Ontology	<i>Initial pilot:</i> Complete. <i>Second phase:</i> Q2 2014
FDA Mini-Sentinel	<i>Goal:</i> Expand medical product safety surveillance capabilities to i2b2 clinical data sources <i>Feedback:</i> Clinical data sources can provide important additional data for medical product safety surveillance. Resources required for data normalization and maintaining additional software at data partners must be carefully considered	<i>Composer:</i> i2b2 <i>Envelope:</i> PopMedNet <i>Processor:</i> i2b2 <i>Other:</i> HQMF	Three-month trial complete
MDPHnet	<i>Goal:</i> Implement Query Envelope security and authentication in existing public health surveillance network <i>Feedback:</i> The Query Envelope enhancements provide standards-based approach for distribution of queries and return of results	<i>Composer:</i> PopMedNet <i>Envelope:</i> PopMedNet <i>Processor:</i> EHR Support for Public Health (ESP) <i>Other:</i> (n/a)	Successful, incorporated into subsequent PopMedNet software releases

Figure 3 The Informatics for Integrating Biology and the Bedside (i2b2) Query Composer. Queries are composed in the graphical i2b2 query builder using the Query Health Consolidated Clinical Document Architecture (C-CDA) data model. These queries are sent to the PopMedNet (PMN) client adapter, which translates the query into a Health Quality Measures Format (HQMF) and sends it to the PopMedNet portal for distribution.



the HQMF translator for translation back into i2b2 format. The Adapter then sends the query to the standard set of i2b2 services for processing. When processing is completed, the Adapter passes the results back to the PopMedNet Data Mart Client. We did not implement QRDA translation in this version of the reference implementation—results are transmitted in i2b2 XML format.

Details on the HQMF translator service have been described elsewhere.⁴² We found that HQMF must be tightly constrained to ensure it represents a functional, automatable query that can be represented within a set of data structures. Implementation guides and harmonization efforts within HL7 will improve this process, but the complexity of existing CQMs might prove computationally problematic.

The HQMF translator relies on an implementation of the Query Health data ontology, which we implemented using the i2b2 ontology cell. This makes extensive use of ‘modifiers’, which were introduced to i2b2 in 2012. They allow additional information about core data elements, such as medication route and dose, by storing attributes about individual patient observations. We

found that they are a very powerful and efficient way of adding metadata to observations in star schema databases.

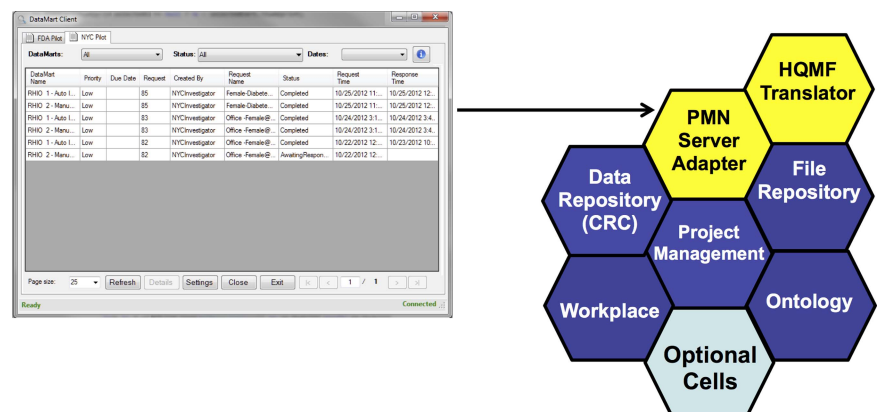
hQuery

An alternative data source for query processing was also selected: the hQuery Gateway, developed by MITRE. hQuery is a document database that uses JavaScript-based map-reduce queries to efficiently search Continuity of Care Documents (CCDs).⁴³ hQuery was not used in a formal pilot, but demonstrating that two very different data sources can both process HQMF makes a strong case for the interoperability of HQMF. Funding for hQuery ceased in 2012 and it is no longer officially supported, but Scoop Health is developing an alternate version of hQuery.

We developed two new components for hQuery:

- A module to convert HQMF queries into native hQuery JavaScript queries, able to process both HQMF revision 1 and 2.¹⁸
- Integration with PopMedNet to enable hQuery to accept queries from and return the results to the sender.

Figure 4 The Informatics for Integrating Biology and the Bedside (i2b2) Query Processing Engine. At each data partner using i2b2, the PopMedNet Data Mart Client sends a Health Quality Measures Format (HQMF) query to an i2b2 instance with a PopMedNet (PMN) server adapter, which translates the query into i2b2 format for execution. Results are returned in i2b2 XML format to the Data Mart Client by way of the server.



When a Query Health-enabled hQuery Gateway receives an HQMF query via PopMedNet, it is converted into the native JavaScript query format, executed, and results are returned to PopMedNet. Results also do not use QRDA format in this version; instead they are presented in the native hQuery JSON format.

We experienced similar difficulties with the complexity of HQMF in implementing the translator to JavaScript. However, the native map-reduce-based query format used by hQuery did eventually permit translation of fairly complex queries. The translator was used in development and testing of the new HQMF standard and is now being used at Research And Development (RAND) Corporation to develop an HQMF to SQL translator that will be applicable to multiple relational database systems.

Pilots

New York City Department of Health

The New York City Department of Health and Mental Hygiene (NYCDOHMH) currently has a network of over 650 primary care practices that respond to distributed queries. This enables their teams of quality improvement specialists to provide regular feedback to the practices in order to respond to public health concerns for both chronic and acute conditions.⁷ However, their current system architecture uses proprietary technology that is vendor-specific and not standards-based. Therefore, the Department is adopting the Query Health platform to add additional practices and health information exchange organizations that are using a wide variety of health information systems.

In 2012, the NYCDOHMH conducted a small pilot of Query Health across three eClinicalWorks EHR practices' test systems. In the pilot, a public health investigator would develop a query in a central i2b2 query builder. This was transmitted to the PopMedNet portal, which distributed the query to the three participating partner sites (Data Marts). Once the query reached and was accepted by a Data Mart, a local i2b2 instance executed the query against the i2b2 data repository. This i2b2 data repository contained a Query Health data model populated with EHR test patients. The practice's result was then transmitted back through PopMedNet, which aggregated the results across practices for viewing in the i2b2 query builder.

This pilot found that, like their existing distributed network, Query Health's simple aggregate counts can successfully provide valuable cross-practice insight with the added benefit of being a standards-based technology solution. In October 2013, NYCDOHMH completed a prototype test of a Query Health-based solution for analyzing the data found within the New York statewide health information exchange (SHIN-NY). In this solution, an aggregate patient CCD is retrieved across the participating exchange partners and loaded into the i2b2 C-CDA system for aggregate querying. Full production release of this system is scheduled for early 2014. It is anticipated that it will provide significant insights into the quality of care delivered particularly in the inpatient settings, which will complement the existing primary care setting solution. Using both systems together will also permit comparisons of Query Health's accuracy, speed, stability, and capabilities. It is worth noting that despite the use of standards, we anticipate that significant work reconciling the various implementations of the data elements and value sets across institutions will remain.

FDA

The FDA's Mini-Sentinel project aims to monitor medical product safety using electronic health data.⁴⁴ Mini-Sentinel is using Query

Health standards implemented through PopMedNet for distributed querying and receiving results within a network of over 130 million individuals.⁹ The Query Health pilot investigates adding i2b2 clinical data repositories to the Mini-Sentinel network to expand the medical product safety monitoring capabilities of the network. The FDA Mini-Sentinel Operations Center team at Harvard Pilgrim Health Care Institute (HPHCI) partnered with Beth Israel Deaconess Medical Center (BIDMC) and Lincoln Peak Partners (the PopMedNet technology partner) to implement the pilot. The i2b2 implementation at BIDMC was the target of the query.

The overall architecture of this pilot was very similar to the New York City pilot—using i2b2, HQMF, and PopMedNet. Also, BIDMC used an i2b2 structure compatible with the Query Health data model. Therefore this pilot's implementation of the reference implementation was fairly straightforward. However, BIDMC did not have resources to host the PopMedNet Data Mart Client software locally. Therefore a new network architecture was developed in which the Data Mart Client was installed securely in the cloud rather than behind the firewall at BIDMC, while the i2b2 data remained behind the BIDMC firewall. Queries were delivered to BIDMC via a secure virtual private network tunnel. This architecture avoided the need to install the PopMedNet software within BIDMC, but still gave BIDMC the ability to review queries before execution and review responses before release. Because only aggregate counts passed through the tunnel, BIDMC viewed this architecture as no more of a security risk than if the software were installed behind their firewall. This pilot was time-limited based on the agreement between HPHCI and BIDMC and was shut down after approximately 3 months of running successfully, but was memorialized in a video available on YouTube.⁴⁵ The pilot demonstrated the feasibility of connecting the FDA Mini-Sentinel network to an i2b2 end point, which could be used to expand Mini-Sentinel to include data from the dozens of healthcare centers that already have their clinical data in i2b2 format.

MDPHnet

MDPHnet is an ONC-funded project overseen by the Massachusetts eHealth Institute in collaboration with the Massachusetts Department of Public Health. MDPHnet is a population-based EHR surveillance network targeting a broad array of health indicators across multiple providers and delivery systems. The project integrates two software systems—PopMedNet and EHR Support for Public Health (ESP)⁴⁶—into a single platform (ESPnet) for population-based surveillance using EHR data.⁴⁷ The Query Health platform is enabling secure, standardized queries on the same architecture. This pilot focused on fully implementing the Query Envelope standard, to demonstrate its flexibility and granular security control.

MDPHnet was already being deployed before Query Health was initiated. This pilot's goal was a technical one only: deployment of a Query Health-compatible version of PopMedNet. This pilot was deployed successfully within MDPHnet and is still being used. The Query Envelope and associated query metadata capture developed as part of this pilot is now integrated fully into the PopMedNet software package and is being used by all PopMedNet-based networks.

DISCUSSION

Query Health is a powerful test of the nation's progress toward an LHS. The initiative has been quite successful: Query Health has created a vendor-neutral, standards-based approach for distributed population health queries; we have delivered an open-

source reference implementation with several alternative configurations; we have co-developed a new revision of HQMF; and we have validated our system for three very different use cases. This work demonstrates enthusiastic collaboration for LHS initiatives across several research groups, government entities, and clinical practices.^{15 17 18 45}

Query Health also uncovered a challenge in our electronic healthcare architecture: most clinical systems, unlike those used in the pilots, do not use common data models. To address this concern, the ONC has launched the Data Access Framework (DAF), an initiative to provide standards and implementation guidance for cross-platform normalized data access. DAF is a multilevel initiative that will encompass standards for both intra- and inter-organizational queries on both individuals and populations. This work is a challenging but necessary step before national-scale distributed networks such as Query Health can become viable on a large scale.

In the meantime, components of and lessons learned from the Query Health initiative are being used for new initiatives from the reference implementation team. The NIH Health Care System Research Collaboratory is developing new distributed networks powered by PopMedNet, including a deeper integration of PopMedNet and i2b2. A pilot recently demonstrated translation of an MDPHnet query to execute directly against an i2b2 database. Both PopMedNet and i2b2 will also be used by portions of PCORI's PCORnet. The i2b2 community can continue to expand HQMF support as federal guidelines for its use are finalized. The i2b2 implementation of the Query Health ontology is being leveraged to develop CCD import tools, which will allow i2b2 to be deployed more broadly. The New York City pilot is continuing to integrate regional health information exchanges into their Query Health network.

CONCLUSION

Query Health has created a vendor-neutral, standards-based approach for distributed population health queries. As Query Health evolves, on-the-fly translation between HQMF and local formats may allow interoperability among systems, creating an infrastructure for comprehensive population health queries. Lessons learned from the Query Health experiment are informing the ONC's DAF, which will encourage data availability for future cross-platform use cases. The present reference implementation provides a common set of components for distributed population health queries that has been validated at three sites.

Author affiliations

¹Partners Healthcare, Boston, Massachusetts, USA

²Harvard Medical School, Boston, Massachusetts, USA

³Massachusetts General Hospital, Boston, Massachusetts, USA

⁴New York City Department of Health and Mental Hygiene, Queens, New York, USA

⁵Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA

⁶MITRE CORP, Burlington, Massachusetts, USA

⁷Allscripts, Burlington, Vermont, USA

⁸Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

Correction notice This article has been corrected since it was published Online First. The corresponding author's email address has been corrected.

Acknowledgements Many thanks to the entire Query Health Reference Implementation team for making this work possible. In particular, we thank: Keith Boone at GE Healthcare, whose work with HQMF has been invaluable in helping Query Health succeed; Nageshwara Bashyam, who managed the reference implementation; Bruce Swan at Lincoln Peak Partners, who led the PopMedNet technical team for the i2b2 integrations; Keely Benson at the Massachusetts eHealth Institute; and John Feikema, who led the initiative at the ONC after Richard Elmore transitioned back to Allscripts. We also would like to thank: Doug Fridsma, Chief Science Officer and Director of Science and Technology at the ONC, who gave us

insight into the broad purpose and vision of Query Health; and Bill Simons, whose expertise in developing the SHRINE distributed network greatly informed our PopMedNet-i2b2 integration.

Contributors JGK wrote the majority of the manuscript and also led the integration of i2b2 with Query Health. MDB was the clinical workgroup chair of the Query Health initiative and developed the Query Health ontology and its i2b2 implementation. MDB also led the New York City pilot. JB leads the coordinating center for PopMedNet and is Director of Data Group for the FDA's Mini-Sentinel Project. MH was the hQuery project lead and implemented the integration between hQuery and PopMedNet. RE was the Query Health project lead at ONC during the development of the majority of this work, and he is therefore intimately familiar with what was accomplished. GMW oversees i2b2 at BIDMC and coordinated the work within BIDMC to join the FDA pilot. These authors each contributed to the manuscript as it related to their individual contribution to Query Health. SNM is Director of Research Computing and Information Systems at Partners' Healthcare and is the chief architect of i2b2. He provided guidance on the preparation of this manuscript, as well as the integration of i2b2 with Query Health.

Funding This work was sponsored by NIH U54LM00874, ONC 90TR0001/01, HHSF2232009100061, AHRQ HS19912-01, and the New York State HEAL 17 Grant 1006090831.

Competing interests None.

Ethics approval Beth Israel Deaconess Medical Center, Harvard Pilgrim Healthcare Institute.

Provenance and peer review Not 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 Friedman CP, Wong AK, Blumenthal D. Achieving a Nationwide Learning Health System. *Sci Transl Med* 2010;2:57cm29(1)-29(3).
- 2 Roundtable on Evidence-Based Medicine, Olsen L, Aisner D, McGinnis JM. *The Learning Healthcare System: Workshop Summary (IOM Roundtable on Evidence-Based Medicine)*. The National Academies Press, 2007. http://www.nap.edu/catalog.php?record_id=11903 (accessed 6 Dec 2013).
- 3 Office of the National Coordinator for Health IT. Federal Health Information Technology Strategic Plan 2011-2015. 2011. <http://www.healthit.gov/sites/default/files/utility/final-federal-health-it-strategic-plan-0911.pdf>
- 4 Fridsma D. Join Query Health in Developing National Standards for Population Queries. Sep 23, 2011. <http://www.healthit.gov/buzz-blog/from-the-onc-desk/queryhealth/> (accessed 14 Mar 2012).
- 5 Weber GM. Federated queries of clinical data repositories: the sum of the parts does not equal the whole. *J Am Med Inform Assoc* 2013;20(e1):e155-61.
- 6 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;16:624-30.
- 7 Buck MD, Anane S, Taverna J, et al. The Hub Population Health System: distributed ad hoc queries and alerts. *J Am Med Inform Assoc* 2011;19:e46-50.
- 8 Natter MD, Quan J, Ortiz DM, et al. An i2b2-based, generalizable, open source, self-scaling chronic disease registry. *J Am Med Inform Assoc* 2013;20:172-9.
- 9 Curtis LH, Weiner MG, Boudreau DM, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoevidenc Drug Saf* 2012;21:23-31.
- 10 Mass eHealth Institute. MDPHnet—Distributed Data Analysis. <http://mehi.masstech.org/what-we-do/hie/mdphnet/about> (accessed 7 Jan 2014).
- 11 NIH Collaboratory. NIH Distributed Research Network. <https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx> (accessed 7 Jan 2014).
- 12 Toh S, Platt R, Steiner JF, et al. Comparative-effectiveness research in distributed health data networks. *Clin Pharmacol Ther* 2011;90:883-7.
- 13 Kim KK, Browe DK, Logan HC, et al. Data governance requirements for distributed clinical research networks: triangulating perspectives of diverse stakeholders. *J Am Med Inform Assoc* 2014;21:714-9.
- 14 Nichols G, Desai J, Elston J, et al. Construction of a multisite datalink using electronic health records for the identification, surveillance, prevention, and management of diabetes mellitus: The SUPREME-DM Project. *Prev Chronic Dis* 2012;9:110311.
- 15 Query Health Pilots Team. i2b2 Query Health Reference Implementation Wiki. i2b2 wiki. <https://community.i2b2.org/wiki/display/QueryHealth/Home> (accessed 7 Jan 2014).
- 16 PopMedNet. Distributed research network technologies for population medicine. <http://www.popmednet.org/> (accessed 7 Jan 2014).

- 17 Query Health Reference Implementation. Standards and Interoperability Framework <http://code.google.com/p/query-health/source/browse>
- 18 PopHealth Initiative. Library to convert HQMF to Javascript. GitHub. <https://github.com/pophealth/hqmf2js> (accessed 6 Dec 2013).
- 19 Klann J, Buck M, Brown JS, et al. Query Health: one step toward a learning health system. In: *Proceedings of the AMIA Symposium*. 2013:781–3.
- 20 Buck M, Klann J, Vogel J, et al. *Query Health Across Communities*. HIMSS 2013 Annual Conference Interoperability Showcase, 2013.
- 21 Klann J, Buck M. Query Health across communities: The New York City Department of Health and Mental Hygiene Pilot. In: *Proceedings of the AMIA Symposium*. 2013:785.
- 22 Prestigiacomo J. Asking population health's unanswered questions. *Healthcare Informatics* 2012;29:14–16. <http://www.healthcare-informatics.com/article/asking-population-healths-unanswered-questions> (accessed 8 Mar 2013).
- 23 Buck M, Elmore R. ONC Project Query Health starts testing its analytics standards. *Health Data Manag* 2012. <http://www.healthdatamanagement.com/news/query-health-distributed-queries-data-analytics-analysis-44013-1.html> (accessed 10 Jan 2014).
- 24 Klann J, Murphy S. Query Health and i2b2: enabling standards-based, multiplatform population health queries. In: *Proceedings of the AMIA Symposium*. 2012:1838.
- 25 Query Health Pilots Team. Query Health pilots synopsis. 2012 Jun. http://wiki.siframework.org/file/view/Query+Health+Pilots+Synopsis_v0+1.docx
- 26 Buck M. *Testimony on the role of clinical documentation for population and public health*. Arlington, VA: ONC HIT Policy Committee Meaningful Use Workgroup and Certification and Adoption Workgroup, 2013. <http://www.healthit.gov/sites/default/files/archive/FACA%20Hearings/2013-02-13%20Policy%3A%20Meaningful%20Use%20and%20Certification%20%26%20Adoption%20WGs,%20Clinical%20Documentation%20Hearing/buckhitpolicytestimony20130206.pdf>
- 27 Elmore R. Quality Measures Workgroup Clinical Quality Public Hearing. Health IT Policy Committee and Health IT Standards Committee: Jun 7, 2012. http://www.healthit.gov/sites/default/files/elmore_testimony_httpc_hitsc_060712_0.pdf
- 28 Platt R, Carnahan R. The U.S. Food and Drug Administration's Mini-Sentinel Program. *Pharmacoepidemiol Drug Saf* 2012;21:1–303. doi:10.1002/pds.3230
- 29 Health Level Seven (HL7). Implementation Guide for CDA Release 2: IHE Health Story Consolidation, DSTU Release 1.1. Ann Arbor, MI: 2012 Jul. Report No.: CDAR2_IG_IHE_CONSOL_DSTU_R1.1_2012JUL.
- 30 Health Level Seven (HL7). Quality Reporting Document Architecture—Category I, DSTU Release 2 (QRDA). Ann Arbor, MI: 2012 Jul. Report No.: CDAR2_QRDA_DSTU_R2_2012JUL.
- 31 Health and Human Services. Health information technology: standards, implementation specifications, and certification criteria for electronic health record technology, 2014 Edition; Revisions to the Permanent Certification Program for Health Information Technology. 77 Federal Register 54163:54163–292. Sep 4, 2012. <https://federalregister.gov/a/2012-20982> (accessed 27 Feb 2013).
- 32 Centers for Medicare and Medicaid Services. Electronic specifications for clinical quality measures. 2012 Dec. http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Electronic_Reporting_Spec.html
- 33 National Quality Forum (NQF). Changing healthcare by the numbers. Washington, D.C: 2012 Mar. Report No.: 2012 NQF Report to Congress. http://www.qualityforum.org/News_And_Resources/Report_to_Congress/Report_to_Congress.aspx
- 34 Health Level Seven (HL7). Representation of the Health Quality Measures Format (eMeasure), Release 2. Ann Arbor, MI: 2012 Sep. Report No.: V3_HQMF_R2_DSTU_2012SEP.
- 35 Health Level Seven (HL7). Quality Data Model Based Health Quality Measures Format (eMeasure) Implementation Guide, Release 1 (US Realm)—Based on HL7 HQMF Release 2.0. Ann Arbor, MI: 2012 Sep. Report No.: V3_IG_HQMF_R1_D1_2012SEP.
- 36 Standards & Interoperability (S&I) Framework. Query Envelope Specification. <http://wiki.siframework.org/Query+Health+--+Query+Envelope> (accessed 3 Feb 2014).
- 37 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(6 Suppl):S45–51. http://journals.lww.com/www-medicalcare/Fulltext/2010/06001/Distributed_Health_Data_Networks__A_Practical_and.9.aspx
- 38 Patient Centered Outcomes Research Institute. PCORnet. PCORnet. <http://pcornet.org> (accessed 7 Jan 2014).
- 39 Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc* 2010;17:124–30.
- 40 i2b2. Informatics for Integrating Biology & the Bedside. <https://www.i2b2.org/> (accessed 7 Jan 2014).
- 41 Patient Centered Outcomes Research Institute. Scalable Collaborative Infrastructure for a Learning Healthcare System (SCILHS). PCORnet. <http://pcornet.org/clinical-data-research-networks/cdrn1-harvard-university-scihls/> (accessed 7 Jan 2014).
- 42 Klann JG, Murphy SN. Computing health quality measures using Informatics for Integrating Biology and the Bedside. *J Med Internet Res* 2013;15:e75.
- 43 The MITRE Corporation. Project hQuery. GitHub. <https://github.com/hquery> (accessed 3 Feb 2014).
- 44 Behrman RE, Benner JS, Brown JS, et al. Developing the Sentinel System — a national resource for evidence development. *N Engl J Med* 2011;364:498–9.
- 45 PopMedNet—i2b2 Integration for ONC Query Health Pilot. 2013. http://www.youtube.com/watch?v=sqDAo6E-b1o&feature=youtuve_gdata_player (accessed 7 Jan 2014).
- 46 Lazarus R, Klompas M, Campion FX, et al. Electronic support for public health: validated case finding and reporting for notifiable diseases using electronic medical data. *J Am Med Inform Assoc* 2009;16:18–24.
- 47 Harvard Pilgrim Health Care Institute. ESPnet: EHR Support for Public Health. <http://esphealth.org/ESPnet/images/ESPnet.html> (accessed 7 Jan 2014).