eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

# We've Only Just Begun – Insights from a 25-Year Journey to Accelerate Health Care Transformation through Delivery System Research

Sarah M. Greene, MPH[1], Paul Wallace, MD[2], Andrew F. Nelson, MPH[3]

*Author Affiliations:*
[1] Health Care Systems Research Network
[2] AcademyHealth
[3] HealthPartners Institute

*Corresponding Author:*
Sarah M. Greene, MPH
sarahgreene@hcsrn.org

## Abstract

Even though it is well known that quality, safety, and patient-centeredness of health care can be improved, leveraging the organizational apparatus of a care delivery environment to render improvement in a consistent and comprehensive manner has proven difficult. The Health Care Systems Research Network (HCSRN), which began as the HMO Research Network, emerged from a desire to improve health and study problems in health care in a systematic and collaborative way, spurring the delivery of true evidence-informed medicine. The HCSRN has honed network-wide data resources, a collaborative culture, and shared infrastructure, enabling multicenter health care research that is often more difficult for researchers working in less integrated settings and across organizational boundaries. The HCSRN's 25-year track record confers both an opportunity and obligation to share what we have learned through our research. Considering the quarter-century since the HCSRN was established, we describe three evolving areas—health data, new health care models, and diversified research teams that must be thoughtfully harnessed to realize a transformed health care ecosystem that generates and learns with research.

eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

## Introduction

Optimizing evidence-informed care in a fragmented health care system is an enduring challenge, as landmark reports from the National Academy of Medicine recognized in “To Err Is Human” and “Crossing the Quality Chasm” (1,2). Even though it is well known that quality, safety, and patient-centeredness can be improved, leveraging the organizational apparatus in a health care environment to render improvement in a consistent and comprehensive manner has proven difficult. Hence, provision of health care is often unsystematic, in that the best available evidence about effective care is not routinely applied (3,4).

The HCSRN, which began as the HMO Research Network, emerged from a desire to improve health and study problems in health care in a systematic and collaborative way (5), spurring the delivery of true evidence-informed medicine. Inspired by the application of systems thinking to create learning organizations (6), the HCSRN has honed network-wide data resources, collaborative culture, and shared infrastructure. This enables large-scale health care research that is often more difficult for researchers working in less integrated settings and across organizational boundaries. The HCSRN’s 25-year portfolio of clinical, epidemiological and health services research not only reflects hard work and privilege, it confers both an opportunity and obligation to share what we have learned through our work. Moreover, as the HCSRN has accrued research results, it has developed a technical and relational infrastructure that enable us to *scale and apply* what we learn to our member health systems and beyond.

Hence, this *eGEMS* journal supplement is more than a compendium that has roots in a multisite network. The methods and insights from the HCSRN and its partners demonstrate the value of shared infrastructure and shared culture. Commonalities in data capabilities and the

shared ideal that research is a public good spurred HCSRN's desire and ability to collaborate. Considering the quarter-century since the HCSRN was established, we describe three areas that have evolved, and must continue to evolve in order, to realize a transformed health care ecosystem that generates and learns from research.

## Three Evolving Aspects of Health System Research

**Health data functions as both a cornerstone and a commodity.** The ubiquity of health data is both an asset and a potential liability. Data are proliferating through sources we did not fully imagine even ten years ago, thanks to wearable devices, social media, the "Internet of Things," and of course, electronic health records (EHRs). This has spawned a commensurate interest in using data to predict events and outcomes—cottage industries related to algorithms, analytics, and visualization have emerged. To successfully improve health and health care in our big data era, we need to focus not only on volume, velocity, and variety—the so-called 3V's of big data (7), but also veracity, validity, and value. Veracity asks us to consider whether the data are accurate and reliable. Validity helps us understand whether the data has the power and representativeness to address a given question. And importantly, the data must yield an answer that has some *value* to stakeholders. Studies of post-marketing surveillance exemplify this—the larger the population, the more precisely incident events can be identified. Embedded in each of these attributes is the need to understand data provenance (i.e., source of the data) and context (i.e., under what circumstances were the data collected?). The HCSRN's Virtual Data Warehouse has long functioned as a federated "big data" repository for our participating systems, and since it is a primary research resource for many of our studies, the attention to

metadata and curation is paramount (8,9). As well, several articles[1] in this special issue illustrate the importance of deep familiarity with the source data.

**New health care and coverage models have significant implications for data-driven health services research.** The prevailing mode of care delivery in ambulatory clinics and hospitals is quickly giving way to health care that is dispersed—virtual care, retail clinics, the hospital at home model, and even health advice that leverages artificial intelligence, are four relatively recent developments poised to alter the health care landscape with lasting impact. Moreover, payment models and health care coverage are shifting from paying for volume of care (leading to a propensity for overuse), to paying for value and managing the health of populations through accountable care entities. This shift from traditional health system delivery arrangements offers new opportunities to conduct research on optimal structure and financing of health care that addresses *individual* circumstances (e.g., sociodemographics, economic disparities, social influences, health status, patient preferences), and *organizational* features (where, how, and by whom care is delivered). As well, this shift brings new challenges for data availability and applicability of findings in new and different care settings. In the HCSRN, comparable data (e.g., on cancer screening, medication orders, or an incident diagnosis) are available with relative consistency, facilitating examination of health system-initiated changes at the patient, clinician, or organizational level (10). Along with the publications in this issue on embedding and measuring specific evidence-based interventions[2]; clinical decision support; and complex care management, the HCSRN's current portfolio includes studies of telemedicine, the use of community-based navigators to support disease management, and impact of benefit designs on medication adherence, all of which offer lessons that can be applied in diverse settings.

[1] See, for example, the papers by Sengupta et al., and Yu et al.
[2] See, for example the paper by Lieu et al.

eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

**Diversified research teams will accelerate our ability to translate research into better health and health care**: Two concurrent influences have ushered in new thinking related to the composition of scientific research teams. The first, team science, has been promulgated by the National Cancer Institute as an intentional approach to creating "coordinated teams of investigators with diverse skills and knowledge…for studies of complex problems with multiple causes." (11). The second is the welcome emphasis on patient- and stakeholder-engaged research, accelerated by creation of the Patient-Centered Outcomes Research Institute (PCORI). By creating multidisciplinary teams and involving end-users of evidence (patients, clinicians, and other decision-makers), health care research can increase both its rigor and its relevance. The HCSRN has decades-long collaborations where multidisciplinary teams might include health services researchers, medical anthropologists, data scientists, and behavioral scientists. In the case of our Mental Health Research Network, a patient/stakeholder advisory board was added recently. Several HCSRN sites have also established formal programs related to the learning health system (12,13), entailing close collaboration with system leaders and frontline clinicians. Many of these sites will be training a new generation of delivery system scientists who develop specific core competencies related to the learning health system (14).

**What do these three forces mean for the HCSRN and the research community?**

This is an exhilarating time for health research. Technology and data are enabling momentous discoveries with the potential to change medical care as we know it. Ever larger research datasets help us undertake observational studies with greater precision (15). Pragmatic clinical trials are a newly potent method for understanding how treatments work under real-world conditions. And yet, we have not yet solved for rampant health inequities, clinician burnout, or high health care costs. Research—a long-standing public good—can help abate this asymmetry. To achieve this, we must continue to forge new relationships with clinical leaders

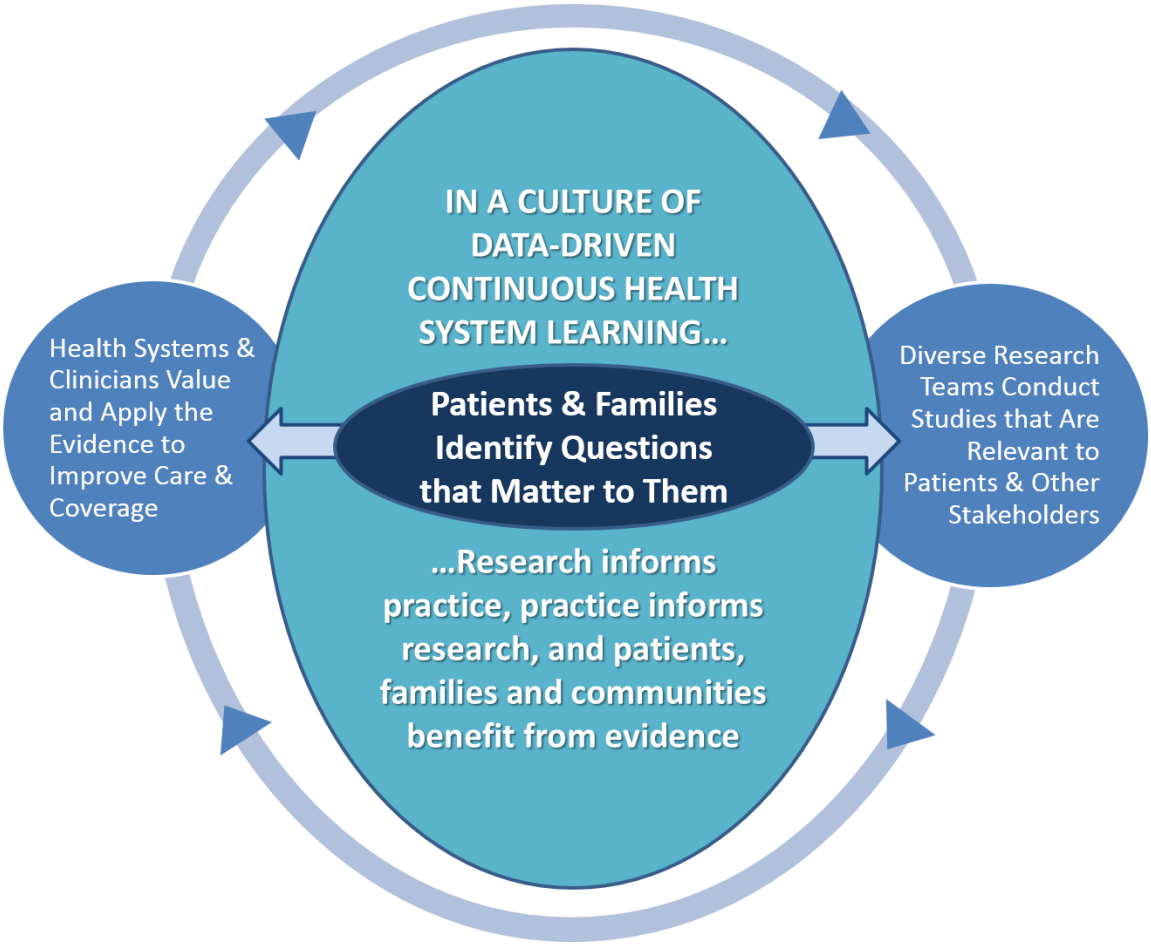
and take down the silos that have slowed the pace of translation. This could mean shifting from a supply-side mindset (conducting research and hoping clinician uptake results) to a demand-driven mindset (undertaking research that is congruent with the urgent needs of health system leaders). The Veterans' Administration Quality Enhancement and Research Initiative (VA QUERI) has pioneered just such a shift, and can serve as an exemplar of this paradigm, purposely linking research activities to clinical care and facilitate adoption of evidence-based care improvements (16-18).

In our experience, the HCSRN has become a useful population laboratory to test interventions in real-world systems—taking action based on an identified care gap or potential health system improvement. And while our member systems are sometimes held up as rarefied settings or "unicorns," where those outside of the system wonder if there are unique facets of [Kaiser Permanente/Geisinger/HealthPartners/etc.] that are not necessarily generalizable to other settings, we have also made progress moving research into practice, which can help inform the research community at large. Health care consultants are often enlisted when health system leaders want focused problem-solving on cost, quality or operational challenges, but what would it take for these leaders to turn to researchers first? We suggest the following steps to encourage closer integration of research and practice, and invigorate collaboration with all stakeholders--researchers, system leaders and clinicians, patients, and communities:

1) Researchers must be conversant in the **evidence base *and* the business case** for their areas of expertise.
2) Researchers must be able to articulate not only the statistical significance, but also the **potential clinical, operational, and practical significance** of our results
3) Researchers must recognize that what we have traditionally valued—publication in high-impact journals and steady funding—may not be as consequential to health care decision-makers or communities, especially relative to care improvement goals. Yet we must be able **convey the value proposition for research** in ways that resonate with system leaders or patient groups.

These steps are achievable if we build on our extant trusting relationships with delivery systems and communities, and share knowledge widely so that it can be applied and amplified. Only then can we overcome the inertia of translation and support patient-centered learning health systems (Figure 1).

*Figure 1. Data-Driven, Patient-Centered Learning Health System*



eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

## The Next 25 Years in Health Care Research

Remarkably, the first publication about the HCSRN in 1998, (when it was still the HMO Research Network) opened with this statement: “Rapid and dramatic change in the health care industry has opened new doors and created expanded opportunities for health services research.” (19) It’s fair to say that this is an evergreen concept—advancements in technology and scientific discovery, with the competing need to control cost and improve quality, guarantee a continuous siege of changes in health care at the individual, population, and organizational level. In recent years, we’ve both witnessed, studied, and led changes in health care brought about by social media, wearables, genomic medicine, telemedicine, immunotherapy, augmented/virtual reality, and consolidation of health and hospital systems, while new “disrupters” aim to diversify options for patients. The potential for innovations is ongoing, but systematic and generalizable investigations into what works best for whom, under what circumstances are no less imperative than they were when medicine was still rudimentary. To the extent the entire research community can leverage a common infrastructure such as the HCSRN, populations will benefit and we’ll see a greater return on our investment in science.

## References

1. Kohn KT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System. Washington, DC: National Academy Press; 1999.

2. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press (US); 2001. PubMed PMID: 25057539.

3. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr. Scientific evidence underlying the ACC/AHA clinical practice guidelines. JAMA. 2009 Feb 25;301(8):831-41. doi:10.1001/jama.2009.205. Erratum in: JAMA. 2009 Apr 15;301(15):1544. PubMed PMID: 19244190.

4. The clinical trials enterprise in the United States: a call for disruptive innovation. In: Califf RM, Filerman GL, Murray RK, Rosenblatt M. Envisioning a transformed clinical trials enterprise in the United States: establishing an agenda for 2020 — workshop summary. Washington, DC: National Academies Press, 2012.

5. Vogt TM, Elston-Lafata J, Tolsma D, Greene SM. The role of research in integrated healthcare systems: the HMO Research Network. Am J Manag Care. 2004 Sep;10(9):643-8. PubMed PMID: 15515997.

6. Senge, Peter M. The Fifth Discipline: the Art and Practice of the Learning Organization. New York: Doubleday/Currency, 1990.

7. Laney D. 3D Data Management: Controlling Data Volume, Velocity and Variety, Gartner Report. (2001). https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf. Accessed January 2019.

8. Hornbrook MC, Hart G, Ellis JL, Bachman DJ, Ansell G, Greene SM, Wagner EH, Pardee R, Schmidt MM, Geiger A, Butani AL, Field T, Fouayzi H, Miroshnik I, Liu L, Diseker R, Wells K, Krajenta R, Lamerato L, Neslund Dudas C. Building a virtual cancer research organization. J Natl Cancer Inst Monogr. 2005;(35):12-25. PubMed PMID: 16287881.

9. Ross TR, Ng D, Brown JS, Pardee R, Hornbrook MC, Hart G, Steiner JF. The HMO Research Network Virtual Data Warehouse: A Public Data Model to Support Collaboration. EGEMS (Wash DC). 2014 Mar 24;2(1):1049. doi:10.13063/2327-9214.1049. eCollection 2014. PubMed PMID: 25848584; PubMed Central PMCID: PMC4371424.

10. Zapka JG, Taplin SH, Solberg LI, Manos MM. A framework for improving the quality of cancer care: the case of breast and cervical cancer screening. Cancer Epidemiol Biomarkers Prev. 2003 Jan;12(1):4-13. Review. PubMed PMID: 12540497.

11. Vogel AL, Hall KL, Fiore SM, Klein JT, Bennett LM, Gadlin H, Stokols D, Nebeling LC, Wuchty S, Patrick K, Spotts EL, Pohl C, Riley WT, Falk-Krzesinski HJ. The Team Science Toolkit: enhancing research collaboration through online knowledge sharing. Am J Prev Med. 2013 Dec;45(6):787-9. doi:10.1016/j.amepre.2013.09.001. PubMed PMID: 24237924.

12. Institute of Medicine (US) Roundtable on Value & Science-Driven Health Care. Learning What Works: Infrastructure Required for Comparative Effectiveness Research: Workshop Summary. Washington: National Academies Press (US); 2011. PubMed PMID: 22013609.

13. Greene SM, Reid RJ, Larson EB. Implementing the learning health system: from concept to action. Ann Intern Med. 2012 Aug 7;157(3):207-10. doi:10.7326/0003-4819-157-3-201208070-00012. PubMed PMID: 22868839.

14. Forrest CB, Chesley FD Jr, Tregear ML, Mistry KB. Development of the Learning Health System Researcher Core Competencies. Health Serv Res. 2017 Aug 4. doi:10.1111/1475-6773.12751. [Epub ahead of print] PubMed PMID: 28777456; PubMed Central PMCID: PMC6051975.

15. Arterburn D, Wellman R, Emiliano A, Smith SR, Odegaard AO, Murali S, Williams N, Coleman KJ, Courcoulas A, Coley RY, Anau J, Pardee R, Toh S, Janning C, Cook A, Sturtevant J, Horgan C, McTigue KM; PCORnet Bariatric Study Collaborative. Comparative Effectiveness and Safety of Bariatric Procedures for Weight Loss: A PCORnet Cohort Study. Ann Intern Med. 2018 Dec 4;169(11):741-750. doi:10.7326/M17-2786. Epub 2018 Oct 30. PubMed PMID: 30383139.

16. Demakis JG, McQueen L, Kizer KW, Feussner JR. Quality Enhancement Research Initiative (QUERI): A collaboration between research and clinical practice. Med Care. 2000 Jun;38(6 Suppl 1):I17-25. Review. PubMed PMID: 10843267.

17. Atkins D, Kilbourne AM, Shulkin D. Moving From Discovery to System-Wide Change: The Role of Research in a Learning Health Care System: Experience from Three Decades of Health Systems Research in the Veterans Health Administration. Annu Rev Public Health. 2017 Mar 20;38:467-487. doi:10.1146/annurev-publhealth-031816-044255. Epub 2017 Jan 11. Review. PubMed PMID: 28125386.

18. Kilbourne AM, Elwy AR, Sales AE, Atkins D. Accelerating Research Impact in a Learning Health Care System: VA's Quality Enhancement Research Initiative in the Choice Act Era. Med Care. 2017 Jul;55 Suppl 7 Suppl 1:S4-S12. doi:10.1097/MLR.0000000000000683. PubMed PMID: 27997456; PubMed Central PMCID:PMC5472006.

19. Durham ML. Partnerships for research among managed care organizations. Health Aff (Millwood). 1998 Jan-Feb;17(1):111-22. PubMed PMID: 9455021.

eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

# HCSRN Special Issue Paper Abstracts by Topic Area



## Effective Approaches to Collaboration

### The Promise of Pragmatic Clinical Trials Embedded in Learning Health Systems

This commentary describes the need for a different context to generate discovery and implement evidence-based advancements to more quickly and efficiently see beneficial effects in health care and health outcomes. Pragmatic clinical trials (PCTs) are a promising type of trial that are conducted within real-world health care delivery systems that embed research like organizations within the Health Care Systems Research Network. A valuable environment for PCTs is learning health systems (LHSs), where clinical practice influences research and vice versa. A goal of LHSs is to operationalize evidence generated by research, particularly PCTs, into improvements that are sustained after a trial ends. PCTs that demonstrate value to health systems and foster implementation could reduce delays in translating research into practice.

### The Healthcare System Research Network (HCSRN) as an environment for Dissemination and Implementation Research: A Case Study of a Multi-Site Research study in Precision Medicine

Context: Variability in the implementation of evidence-based practice into healthcare leads to deficiencies in quality of care across healthcare systems. Dissemination and Implementation research works to bridge this gap. We present a case study of a currently active dissemination and implementation study created through collaborations developed within the Healthcare Systems Research Network (HCSRN).

Case description: The “Implementing Universal Lynch Syndrome Screening (IMPULSS)” study (NIH R01CA211723) involves seven HCSRN healthcare systems and two external healthcare systems. the

IMPULSS study is designed to describe and explain individual organizational variability around Lynch Syndrome screening to identify minimally necessary and sufficient factors in different organizational contexts important for implementing LS screening programs, and will create a toolkit to facilitate organizational decision making around implementation and improvement of such programs in healthcare systems.

Major Themes: The strengths of the HCSRN in the areas of 1) a culture of collaboration, 2) standardization of data and processes across systems, and 3) researchers embedded in diverse healthcare systems are what make the IMPULSS study possible. We describe how these major themes contribute to the IMPULSS study specifically and create an environment for implementation studies within the HCSRN.

Conclusion: Given the importance of conducting research in real world settings to improve patient outcomes, the unique strengths of the HCSRN are of vital importance. The IMPULSS study is one case example of how the strengths of the HCSRN make it an excellent environment for research on implementing evidence based practices in healthcare systems.

### Collaborating on Data, Science, and Infrastructure: The 20-Year Journey of the Cancer Research Network

The Cancer Research Network (CRN) is a consortium of 12 research groups, each affiliated with a nonprofit integrated healthcare delivery system, that was first funded in 1998. The overall goal of the CRN is to support and facilitate collaborative cancer research within its component delivery systems. This paper describes the CRN's 20-year experience and evolution. The network combined its members' scientific capabilities and data resources to create an infrastructure that has ultimately supported over 275 projects. Insights about the strengths and limitations of electronic health data for research, approaches to optimizing multidisciplinary collaboration, and the role of a health services research infrastructure to complement traditional clinical trials and large observational datasets are described, along with recommendations for other research consortia.

## Optimizing Source Data from EHRs to Support Rigorous Research

### Improving Health Care with Advanced Analytics: Practical Considerations

Artificial intelligence is becoming ubiquitous in healthcare, largely through machine learning and predictive analytics applications. Recent applications of AI to common health care scenarios, such as screening and diagnosing, has fueled optimism about the use of advanced analytics to improve care. Careful and objective considerations need to be made before implementing an advanced analytics solution. Critical evaluation before, during, and after its implementation will ensure safe care, good

outcomes, and the elimination of waste. In this commentary we offer basic practical considerations for developing, implementing, and evaluating such solutions based on many years of experience.

## Learning to Share Healthcare Data

Since 2001 there has been a steady stream of work on Common Data Models being done by different organizations. Each has its own needs and directives, yet the core subject areas remain remarkably similar. The timeline and core entities of these models create a compelling picture of this work, which is crucial to research and improving our healthcare delivery. Work continues on these models, keeping pace with innovations that come from advances in genetics, pharmacology and technology.

## Identification of Incident Uterine Fibroids Using Electronic Medical Record Data

Background: Uterine fibroids are the most common benign tumors of the uterus and are associated with considerable morbidity. Diagnosis codes have been used to identify fibroid cases but their accuracy, especially for incident cases, is uncertain.

Methods: We performed medical record review on a random sample of 617 women who received a fibroid diagnosis during 2012-2014 to assess diagnostic accuracy for incident fibroids. We developed 2 algorithms aimed at improving incident case-finding using classification and regression tree analysis that incorporated additional electronic healthcare data on demographics, symptoms, treatment, imaging, healthcare utilization, comorbidities and medication. Algorithm performance was assessed using medical record as gold standard.

Results: Medical record review confirmed 482 fibroid cases as incident, resulting a 78% positive predictive value (PPV) for incident cases based on diagnosis codes alone. Incorporating additional electronic data, the first algorithm classified 395 women with a pelvic ultrasound on diagnosis date but none before as incident cases. Of these, 344 were correctly classified, yielding an 87% PPV, 71% sensitivity, and 62% specificity. A second algorithm built on the first algorithm and further classified women based on a fibroid diagnosis code of 218.9 in 2 years after incident diagnosis and lower body mass index; yielded 93% PPV, 53% sensitivity, and 85% specificity.

Conclusions: Compared to diagnosis codes alone, our algorithms using fibroid diagnosis codes and additional electronic data improved identification of incident cases with higher PPV, and high sensitivity or specificity to meet different aims of future studies seeking to identify incident fibroids from electronic data.

eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

## Impact of ICD-10-CM Transition on Mental Health Diagnoses Recording

Objective: This study examines the impact of the transition from ICD-9-CM to ICD-10-CM diagnosis coding on the recording of mental health disorders in Electronic Health Records (EHR) and claims data in ten large health systems. We present rates of these diagnoses across two years spanning the October 2015 transition.

Methods: Mental health diagnoses were identified from claims and EHR data at ten healthcare systems in the Mental Health Research Network (MHRN). Corresponding ICD-9-CM and ICD-10-CM codes were compiled and monthly rates of people receiving these diagnoses were calculated for one year before and after the coding transition.

Results: For seven of eight diagnostic categories examined, monthly were comparable during the year before and the year after the ICD-10-CM transition. In the remaining category, psychosis excluding schizophrenia spectrum disorders, aggregate monthly rates of decreased markedly with the ICD-10-CM transition, from 48 to 33 per 100,000. We propose that the change is due to features of General Equivalence Mappings (GEMS) embedded in the EHR.

Conclusions: For most mental health conditions, the transition to ICD-10-CM appears to have had minimal impact. The decrease seen for psychosis diagnoses in these health systems is likely due to changes associated with EHR implementation of ICD-10-CM coding rather than an actual change in disease prevalence. It is important to consider the impact of the ICD-10-CM transition for all diagnostic criteria used in research studies, quality measurement, and financial analysis during this interval.

## Data Quality Assessment and Multi-Organizational Reporting: Tools to Enhance Network Knowledge

Objective: Multi-organizational research requires a multi-organizational data quality assessment (DQA) process that combines and compares data across participating organizations. We demonstrate how such a DQA approach complements traditional checks of internal reliability and validity by allowing for assessments of data consistency and the evaluation of data patterns in the absence of an external “gold standard”.

Methods: We describe the DQA process employed by the Data Coordinating Center (DCC) for Kaiser Permanente’s (KP) Center for Effectiveness and Safety Research (CESR). We emphasize the CESR DQA reporting system that compares data summaries from the eight KP organizations in a consistent, standardized manner.

Results: We provide examples of multi-organization comparisons from DQA to confirm expectations about different aspects of data quality. These include 1) comparison of direct data extraction from the electronic health records (EHR) 2) comparison of non-EHR data from disparate sources.

Discussion: The CESR DCC has developed codes and procedures for efficiently implementing and reporting DQA. The CESR DCC approach is to 1) distribute DQA tools to empower data managers at each organization to assess their data quality at any time 2) summarize and disseminate findings to address data shortfalls or document idiosyncrasies and 3) engage data managers and end-users in an exchange of knowledge about the quality and its fitness for use.

Conclusion: KP CESR DQA model is applicable to networks hoping to improve data quality. The multi-organizational reporting systems promotes transparency of DQA, adds to network knowledge about data quality, and informs research.

## Comparing Prescribing and Dispensing of the PCORnet Common Data Model using PCORnet Antibiotics and Childhood Growth Study

Researchers often use prescribing data from electronic health records (EHR) or dispensing data from medication or medical claims to determine medication utilization. However, neither source has complete information on medication use. We compared antibiotic prescribing and dispensing records for 200,395 patients in the National Patient-Centered Clinical Research Network (PCORnet) Antibiotics and Childhood Growth Study. We stratified analyses by delivery system type [closed integrated (cIDS) and non-cIDS]; 90.5% and 39.4% of prescribing records had matching dispensing records, and 92.7 % and 64.0% of dispensing records had matching prescribing records at cIDS and non-cIDS, respectively. Most of the dispensings without a matching prescription did not have same-day encounters in the EHR, suggesting they were medications given outside the institution providing data, such as those from urgent care or retail clinics. The sensitivity of prescriptions in the EHR, using dispensings as a gold standard, was 99.1% and 89.9% for cIDS and non-cIDS, respectively. Only 0.7% and 6.1% of patients at cIDS and non-cIDS, respectively, were classified as false-negative, i.e. entirely unexposed to antibiotics when they in fact had dispensings. These patients were more likely to have a complex chronic condition or asthma. Overall, prescription records worked well to identify exposure to antibiotics. EHR data, such as the data available in PCORnet, is a unique and vital resource for clinical research. Closing data gaps by understanding why prescriptions may not be captured can improve this type of data, making it more robust for observational research.

## Assessing and Minimizing Re-identification Risk in Research Data Derived from Healthcare Records

Background – Sharing of research data derived from health system records supports the rigor and reproducibility of primary research and can accelerate research progress through secondary use. But public sharing of such data can create risk of re-identifying individuals, exposing sensitive health information.

Method – We describe a framework for assessing re-identification risk that includes: identifying data elements in a research dataset that overlap with external data sources, identifying small classes of

records defined by unique combinations of those data elements, and considering the pattern of population overlap between the research dataset and an external source. We also describe alternative strategies for mitigating risk when the external data source can or cannot be directly examined.

Results – We illustrate this framework using the example of a large database used to develop and validate models predicting suicidal behavior after an outpatient visit. We identify elements in the research dataset that might create risk and propose a specific risk mitigation strategy: deleting indicators for health system (a proxy for state of residence) and visit year.

Discussion – Researchers holding health system data must balance the public health value of data sharing against the duty to protect the privacy of health system members. Specific steps can provide a useful estimate of re-identification risk and point to effective risk mitigation strategies.

## Applying Electronic Data for Point-of-Care Improvement

### Developing a Prognostic Information System for Personalized Care in Real Time

Context. Electronic medical records hold promise to transform clinical practice. However, technological and other barriers may preclude using them to guide care in real time. We used the Virtual Data Warehouse (VDW) to develop a tool that enables physicians to generate real-time, personalized prognostic information about survival after cancer.

Case description. Patients with cancer often ask their oncologists, “Have you ever seen a patient like me?” To help oncologists answer this question, we developed a prototype Prognostic Information System (PRISM), a web-based tool that gathers data about the index patient from Kaiser Permanente’s clinical information systems, selects a historical cohort of similar patients, and displays the survival curve of the similar patients relative to key points in their treatment course.

Findings and major themes. The prototype was developed by a multidisciplinary team with expertise in oncology, research, and technology. We have completed two rounds of user testing and refinement. Successful development rested on: (1) executive support and a clinical champion; (2) collaboration among experts from multiple disciplines; (3) starting with simple cases rather than ambitious ones; (4) extensive research experience with the Virtual Data Warehouse, related databases, and an existing query tool; (5) following agile software development principles, especially iterative user testing.

Conclusion. Clinical data stored in health care systems’ electronic medical records can be used to personalize clinical care in real time. Development of prognostic information systems can be accelerated by collaborations among researchers, technology specialists, and clinicians and by use of existing technology like the Virtual Data Warehouse.

eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

## Development of a Clinical Decision Support System for Pediatric Abdominal Pain in Emergency Room Settings Across Two Health Systems Within the HCSRN

Background: Appendicitis is a common surgical emergency in children, yet the diagnosis can be challenging. An electronic health record (EHR) - based clinical decision support (CDS) system called Appy CDS was designed to help guide management of pediatric patients with acute abdominal pain.

Methods: This prospective pragmatic cluster-randomized clinical trial in 17 emergency departments of two large HCSRN-affiliated care systems began in October 2016 and runs through July 2019. Appy CDS was built independently but synergistically at each site using well-established platforms.

Results: Although differences across sites in terms of design and implementation, using simple screening questions and automated exclusions, we can identify a population where use of Appy CDS is indicated, with an appendicitis rate of 11.3%, consistent across study arms and sites.

Discussion: These 2 HCSRN sites designed the intervention to capture the population at risk for appendicitis and deliver CDS to that population while remaining locally relevant and adhering to organizational preferences. Despite different approaches to point-of-care CDS, the interventions have enrolled similar populations with nearly identical background rates of appendicitis.

Next Steps: The Appy CDS tool is testing a clinical prediction algorithm to reduce reliance on imaging in the pediatric population. The project team will continue to address issues around alert fatigue and provide feedback to sites for the remainder of the study period. These sophisticated web-based EHR-linked CDS systems provide a personalized risk assessment and tailored recommendations at the point of care. These novel approaches could serve as the basis for future ED interventions.

## Challenges of population-based measurement of suicide prevention activities across multiple health systems

Suicide is a preventable public health problem. Zero Suicide (ZS) is a suicide prevention framework currently being evaluated by Mental Health Research Network investigators embedded in six Health Care Systems Research Network (HCSRN) member health systems implementing ZS. This paper describes ongoing collaboration to develop population-based process improvement metrics for use in, and comparison across, these and other health systems. Robust process improvement metrics are sorely needed by the hundreds of health systems across the country preparing to implement their own best practices in suicide care. Here we articulate three examples of challenges in using health system data to assess suicide prevention activities, each in ascending order of complexity: 1) Mapping and reconciling different versions of suicide risk assessment instruments across health systems; 2) Deciding what should count as adequate suicide prevention follow up care and how to count it in different health systems with different care processes; and 3) Trying to determine whether a safety planning discussion took place between a clinician and a patient, and if so, what actually happened. To develop broadly applicable metrics, we have advocated for standardization of care processes and their documentation,

encouraged standardized screening tools and urged they be recorded as discrete EHR variables, and engaged with our clinical partners and health system data architects to identify all relevant care processes and the ways they are recorded in the EHR so we are not systematically missing important data. Serving as embedded research partners in our local ZS implementation teams has facilitated this work.

## Priorities Wizard: multisite web-based primary care chronic disease clinical decision support improved clinical outcomes with high use rates and high clinician satisfaction rates

Abstract: Priorities Wizard is an electronic health record (EHR)-linked, web-based clinical decision support (CDS) system designed and implemented at multiple HCSRN sites to support high quality outpatient chronic disease and preventive care. The CDS system (a) identifies patients who could substantially benefit from evidence-based actions; (b) presents prioritized evidence-based treatment options to both patient and clinician at the point of care; and (c) facilitates efficient ordering of recommended medications, referrals or procedures.

Methods: The CDS system extracts relevant data from electronic health records (EHRs), processes the data using Web-based clinical decision support algorithms, and displays the CDS output seamlessly on the EHR screen for use by the clinician and patient. Through a series of National Institutes of Health (NIH)-funded projects led by HealthPartners Institute and the HealthPartners Center for Chronic Care Innovation and HCSRN partners, Priorities Wizard has been expanded to include 14 clinical domains.

Results: Cluster-randomized trials show that this CDS system significantly improved glucose and blood pressure control in diabetes patients, reduced 10-year CV risk in high-CV risk adults without diabetes, improved management of smoking in dental patients, and improved high blood pressure identification and management in adolescents.1-4 CDS is used at 71-77% of targeted visits, 85-98% percent of clinicians are satisfied with the CDS system, and 94% report they would recommend it to colleagues.

Conclusions: Recently developed EHR-linked, Web-based CDS systems have significantly improved chronic disease care outcomes and have high use rates and primary care clinician satisfaction.

## Using Self-Reported Data to Segment Older Adult Populations with Complex Care Needs

Background: Tailored care management requires effectively segmenting heterogeneous populations into actionable subgroups. Using patient reported data may help identify groups with care needs not revealed in traditional clinical data.

Methods: We conducted retrospective segmentation analyses of 9,617 Kaiser Permanente Colorado members age 65 or older at risk for high utilization due to advanced illness and geriatric issues who had

eGEMs The Journal for Electronic Health Data and Methods

**PREPUBLICATION COPY**
Forthcoming in *eGEMs*
Vol. 7, Iss. 1, April 2019

completed a Medicare Health Risk Assessment (HRA) between 2014 and 2017. We separately applied clustering methods and latent class analyses (LCA) to HRA variables to identify groups of individuals with actionable profiles that may inform care management. HRA variables reflected self-reported quality of life, mood, activities of daily living (ADL), urinary incontinence, falls, living situation, isolation, financial constraints, and advance directives. We described groups by demographic, utilization, and clinical characteristics.

Results: Cluster analyses produced a 14-cluster solution and LCA produced an 8-class solution reflecting groups with identifiable care needs. Example groups included: frail individuals with memory impairment less likely to live independently, those with poor physical and mental well-being and ADL limitations, those with ADL limitations but good mental and physical well-being, and those with few health or other limitations differentiated by age, presence or absence of a documented advance directive, and tobacco use.

Conclusions: Segmenting populations with complex care needs into meaningful subgroups can inform tailored care management. We found groups produced through cluster methods to be more intuitive, but both methods produced actionable information. Applying these methods to patient-reported data may make care more efficient and patient-centered.