

## TECHNICAL REPORT

# Data and knowledge standards for learning health: A population management example using chronic kidney disease

Blake Cameron<sup>1</sup>  | Brian Douthit<sup>2</sup> | Rachel Richesson<sup>2</sup>

<sup>1</sup>Division of Nephrology, Department of Medicine, Duke University School of Medicine, Durham, North Carolina

<sup>2</sup>Duke University School of Nursing, Durham, North Carolina

**Correspondence**

Blake Cameron, Division of Nephrology, Department of Medicine, Duke University School of Medicine, Durham, NC 27705.  
Email: blake.cameron@duke.edu

**Abstract**

The widespread creation of learning health care systems (LHSs) will depend upon the use of standards for data and knowledge representation. Standards can facilitate the reuse of approaches for the identification of patient cohorts and the implementation of interventions. Standards also support rapid evaluation and dissemination across organizations. Building upon widely-used models for process improvement, we identify specific LHS activities that will require data and knowledge standards. Using chronic kidney disease (CKD) as an example, we highlight the specific data and knowledge requirements for a disease-specific LHS cycle, and subsequently identify areas where standards specifications, clarification, and tools are needed. The current data standards for CKD population management recommendations were found to be partially ambiguous, leading to barriers in phenotyping, risk identification, patient-centered clinical decision support, patient education needs, and care planning. Robust tools are needed to effectively identify patient health care needs and preferences and to measure outcomes that accurately depict the multiple facets of CKD. This example presents an approach for defining the specific data and knowledge representation standards required to implement condition-specific population health management programs. These standards specifications can be promoted by disease advocacy and professional societies to enable the widespread design, implementation, and evaluation of evidence-based health interventions, and the subsequent dissemination of experience in different settings and populations.

**KEYWORDS**

chronic kidney disease, learning health, population health, standards

## 1 | INTRODUCTION AND BACKGROUND

With the rapid transition to value-based payment mechanisms, health care organizations in the United States must learn to efficiently manage the health of local populations to which they are accountable.<sup>1</sup> Managing the cost and quality of care at a population level requires adoption of learning health principles. Specifically,

organizations must apply technologies and develop organizational processes that support an iterative model of improvement. A growing body of literature documents the barriers to adopting these practices, arising from overwhelming system complexity, suboptimal data quality and analytical tools, and misaligned health care policies.<sup>2-6</sup> In particular, health care organizations struggle to leverage data stored within proprietary electronic health record (EHR) systems to gain

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2018 The Authors. *Learning Health Systems* published by Wiley Periodicals, Inc. on behalf of the University of Michigan

meaningful insights into processes of care and to test and evaluate clinical interventions. Lack of standardization hampers the diffusion of evidence and best practices across organizations.

Learning health systems (LHS), which unify clinical care and research in a cycle of continuous improvement, will depend upon the use of standards to drive improvements in care at a national scale. Such standards take many forms, spanning the entire hierarchy of organization, from low-level data elements to high-level knowledge structures. Data standards may specify variable definitions and data types, data elements, or value sets; or may encompass entire coding or terminological systems such as International Classification of Diseases (ICD-10) or the Systematized Nomenclature of Medicine (SNOMED CT), Logical Observation Identifiers Names and Codes (LOINC), and RxNorm. Standards also include specifications for representing knowledge (eg, the Health Level 7 [HL7] Clinical Document Architecture [CDA] specification for a patient care plan), for exchanging of health information (HL7 v2 or Fast Healthcare Interoperability Resources [FHIR]), and for creating domain-specific “functional profiles” in the EHR (eg, functional profiles specific to behavioral health practice<sup>7</sup>). In an effort to rationalize the growing volume and complexity of standards, the Office of the National Coordinator for Health Information Technology (ONC) maintains a clearinghouse for the identification, assessment, and documentation of myriad health care standards and specifications.<sup>8</sup> Since 2015, the ONC has released the Interoperability Standards Advisory (ISA) reference document that provides an inventory of best available standards to address specific interoperability needs. The ISA is open for public comment and updated each year.

Though not data standards per se, computable phenotype definitions (EHR search algorithms that identify patient cohorts according to clinical features or disease status) depend upon standards to aggregate data from multiple sources and data types. Explicit, standardized phenotype definitions facilitate the reuse of clinical data and tools for population management, quality measurement, and research.<sup>9</sup> Further, the computable representation of knowledge—for example, knowledge embedded in clinical practice guidelines or drug-interaction databases—can support the dissemination and rapid adoption of automated clinical decision support tools, such as alerts, reminders, and customized order sets.<sup>10</sup>

In this paper, we use existing models for evidence-based medicine (EBM), learning health, and process improvement to identify data-dependent activities, which by definition will benefit from the use of standards. We identify relevant data domains for each data-dependent activity, and look at national standards to evaluate the clarity and specificity of those current data standards in the context of a population management program for CKD.

## 2 | DATA-DEPENDENT ACTIVITIES AND STANDARDS FOR LEARNING HEALTH SYSTEMS

In an early roundtable on learning health, the Institute of Medicine set a goal that “by the year 2020, ninety percent of clinical decisions will be supported by accurate, timely, and up-to-date information, and will reflect the best available evidence.”<sup>11</sup> The PICO (Patient, Intervention,

Comparison, Outcome) mnemonic aids clinicians in framing clinical questions to locate the best available evidence in the medical literature. When considered through the lens of learning health care and medical decision-making, the PICO model reveals important data requirements. These include the digital capture and grouping of individual health characteristics (*the patient*), the identification of potential evidence-based clinical *interventions*, the health care services already received by similar patients (*the comparison*), and the aggregation of clinical and patient-reported data across multiple sources (*the outcome*). Each of these functions depends on data standards that harmoniously represent patient features, health care activities, and best available evidence.

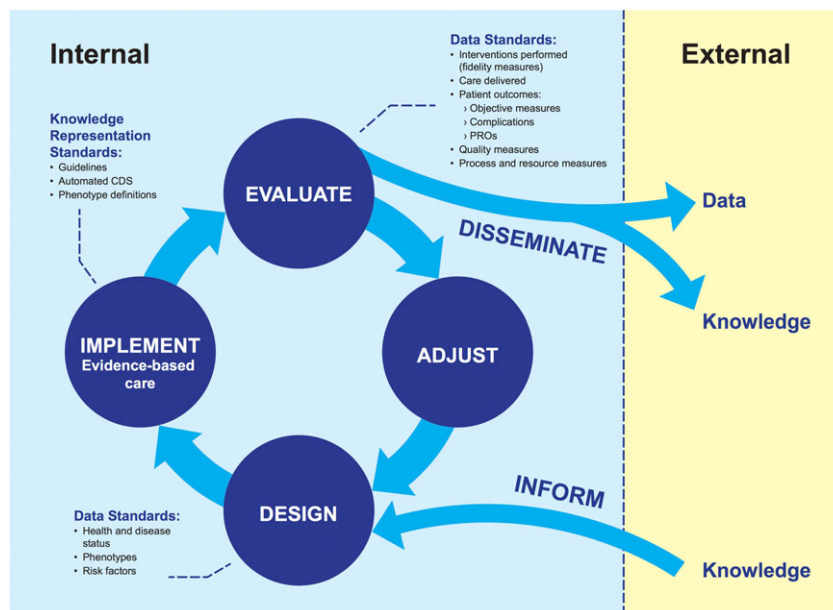
Greene et al<sup>12</sup> proposed a pragmatic framework for LHS that couples quality improvement with the need to apply both local experience and external evidence in designing improvement programs. The exchange of internal and external knowledge of patient features, health care activities, and medical evidence requires a shared nomenclature. These activities cannot be scaled across complex information systems without robust data and knowledge representation standards.

Building upon a generalizable framework for LHS activities, we can identify the different types of data and corresponding standards required to develop informatics processes and tools to support specific clinical problems (Figure 1). Shared standards are essential to identify relevant patient cohorts, to enable clinical decision support applications, and to evaluate the impact of interventions across multiple organizations. When multiple parties agree upon standards, they can exchange data or software applications developed to work with those standards, bringing us closer to the ideal of a national learning health network. Through bottom-up improvement efforts and clearinghouses like the ISA, standards can be assessed, improved, and promoted to facilitate broad reuse and scaling.

The data-dependent LHS activities (design, implement, evaluate, and adjust) are presented above with a number of generalizable data types and knowledge artifacts that can be used to support EBM and evidence generation for a specific condition. As presented, this (and other LHS models) do not specify value sets and configurations of standards. To our knowledge, there is no commonly agreed upon approach to identifying and assessing the fitness of data and knowledge representation standards for a defined condition or for condition-specific population management programs. Because data and knowledge representation standards can enable the sharing of tools and knowledge dissemination from internal LHS to others, the identification and adoption of standards that are relevant to a given condition will ultimately improve the health of affected populations on a national scale. Using chronic kidney disease as an example, the next section explores how specific data standards relevant to a specific condition could be used to advance national population health goals and population health.

## 3 | STANDARDS-ENABLED LEARNING HEALTH SYSTEMS—THE EXAMPLE OF CHRONIC KIDNEY DISEASE

Chronic kidney disease (CKD) affects an estimated 40 million U.S. adults.<sup>13</sup> It is an illness defined by objective criteria, and as such, is a



**FIGURE 1** Data-dependent activities for learning health care system and types of data and knowledge for each. CDS, clinical decision support

model condition for illustrating the centrality of data standards in a rapid learning health system.<sup>14</sup> As a leading cause of catastrophic disability, CKD has profound implications not only for patients but also for organizations operating under accountable care arrangements. Each year, about 150 000 patients with CKD progress to complete kidney failure, known as end-stage renal disease (ESRD), requiring kidney transplantation or life-long dialysis treatments to survive. ESRD causes significant morbidity and mortality, with a life expectancy on dialysis of only 5 years.<sup>15</sup> Health care costs for patients with CKD rise in lockstep with the severity of illness, evidenced by the staggering cost of caring for the 600 000 patients with ESRD on dialysis; though ESRD patients make up only 1% of the Medicare population, they account for nearly 7% of all Medicare spending.<sup>16</sup> Prompt detection and treatment of CKD can slow or halt progression to ESRD and prevent complications. Yet CKD is often under-recognized and under-treated in usual care, leading to avoidable morbidity, mortality, and cost.

To reduce the burden and complications of CKD nationally, Mendu and colleagues<sup>17</sup> recently proposed a population health management framework that includes recommendations for the identification of the illness, the management of various risks and complications, and a set of quality measures. A summary of key elements for the ideal care of CKD and recommended organizational and EHR system initiatives for each phase of care are presented in Table 1. These elements provide disease-specific examples of a set of generalizable data requirements for various learning health activities, such as disease identification, adherence to clinical practice guidelines, automated decision support.

## 4 | DATA STANDARDS FOR CKD POPULATION HEALTH MANAGEMENT

Using Mendu's<sup>17</sup> conceptual framework for CKD care as an example of a disease-specific population health management program that a

LHS might undertake, we systematically identified the data-dependent activities and the types of data standards required for each. We summarized the relevant standards by activity, assessed the maturity and fitness of existing standards for representing data relevant to CKD population management, and identified areas where standards were missing or needing clarification.

### 4.1 | Patient and risk factor identification using computable phenotypes

As a foundational capability, LHS must reliably identify populations at risk of clinical deterioration or for generating excess cost. Computable phenotype definitions are structured EHR database queries that identify cohorts of patients with certain diseases or clinical profiles for a variety of purposes, including disease management registries and quality improvement programs. Numerous studies examining the accuracy of diagnostic coding have shown that encounter-level diagnosis codes alone are insufficient to identify the full population of individuals with a particular condition.<sup>18,19</sup> Consequently, computable phenotype definitions are typically constructed using a combination of data elements and logic that include diagnosis codes from encounters or EHR problem lists, clinical documentation, diagnostic test results, prescribed therapies including medications, and repeated measures or time. Each of these categories of data have associated standards, such as LOINC (labs), RxNorm (medications), and ICD-10-CM (diagnoses).

The ease with which a particular disease cohort can be identified depends in part on the nature of the illness and the availability of relevant structured data in the EHR. In CKD, unlike many other conditions, the illness can be identified by abnormalities in recent laboratory tests, without relying upon clinicians' subjective documentation or diagnostic coding choices. Professional society guidelines define CKD as a (1) a sustained impairment in kidney function (estimated glomerular filtration rate  $<60$  mL/min/1.73m<sup>2</sup> for at least 3 months); and/or (2) sustained evidence of kidney damage

**TABLE 1** Mendu et al framework for CKD population health management<sup>17</sup>

Stage Within Care Spectrum	Key Elements for Ideal Care	Proposed Initiative
CKD recognition (based on eGFR, proteinuria)	<ol style="list-style-type: none"> <li>1. CKD patients identified by EHR</li> <li>2. Early treatment initiated through electronic consultation</li> </ol>	EHR-based referral alert/ guidelines; electronic specialist consultation (e-consult)
Stemming CKD progression (ie, ACE inhibitor/ARB use, blood pressure/ diabetes control, nephrotoxin avoidance)	<ol style="list-style-type: none"> <li>1. Adherence to clinical guidelines</li> <li>2. Collaborative care plan eliminating redundant visits for routine care, improving appropriate referral</li> </ol>	EHR-based CKD decision support; CKD patient registry and quality reporting; EHR-based implementation of PC/renal CCA
Treating complications (ie, cardiovascular risk, anemia, bone mineral disease)	<ol style="list-style-type: none"> <li>1. Adherence to clinical guidelines</li> <li>2. Close nonvisit collaboration with PCPs through collaborative care</li> </ol>	EHR-based CKD quality metrics; CKD registry; EHR-based implementation of PC/renal CCA
Renal replacement preparation for advanced CKD (ie, education, access evaluation, transplant evaluation)	<ol style="list-style-type: none"> <li>1. Early referral to nephrology with EHR-based decision support with referral guidelines, electronic consultation</li> <li>2. CKD registry used to guide prediction of progression</li> <li>3. Collaborative care with vascular surgery and transplantation with renal to facilitate coordinated care</li> <li>4. Multidisciplinary clinic to facilitate timely initiation of outpatient renal replacement, increased PD use, transplantation evaluation</li> </ol>	EHR-based referral alert/ guidelines; e-consult; CKD registry to guide prediction of progression; CCA with vascular surgery and renal transplantation; checklist for RRT initiation; multidisciplinary clinic to facilitate advanced CKD care
ESRD patient management (ie, reducing readmissions, proactive access management)	<ol style="list-style-type: none"> <li>1. Care coordinators to reduce readmissions, coordinate care</li> <li>2. Home care management</li> <li>3. Establishing network of dialysis clinics for data sharing, focused on reducing readmissions</li> </ol>	Home-based care; care coordinator (care coordinators specialized in ESRD); dialysis clinic collaborative
Throughout care spectrum	Patient engagement to facilitate adherence and understanding of disease process	Shared decision making; Vidscrips; patient-reported outcome measures

Reproduced with permission.

Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; CCA, collaborative care agreement; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; EHR, electronic health record; ESRD, end-stage renal disease; PC, primary care; PCPs, primary care providers; PD, peritoneal dialysis; RRT, renal replacement therapy.

(urinary albumin excretion  $\geq 30$  mg/24 h).<sup>20</sup> CKD is further staged according to the severity of abnormalities in kidney function or albuminuria.<sup>20</sup> This laboratory data is frequently obtained in the course of routine medical care and is represented in a structured form within the EHR, making a computable phenotype possible.

While laboratory-based identification is one method, multiple guideline-derived computable phenotype definitions for CKD have been proposed and are being used in practice. Among the earliest published definitions is one from the Cleveland Clinic CKD registry, which classifies CKD status according to laboratory indicators of persistently reduced kidney function and identified comorbidities, but differs from current practice guidelines in that it does not take into account laboratory markers of kidney damage.<sup>21</sup> The Centers for Medicare and Medicaid Services Chronic Conditions Warehouse definition of CKD uses only diagnosis codes, and includes diagnosis codes for a number of conditions that are not considered to be CKD by clinical practice guidelines.<sup>22</sup> A recent systematic comparison of these CKD phenotype definitions and others found poor agreement among definitions, with cohorts returned differing substantially.<sup>23</sup> This issue is highlighted by McDonald and Overhage,<sup>24</sup> where researchers may identify CKD populations from the EHR based on coded diagnoses, or on laboratory values. Challenges exist in both, as the diagnosis is dependent on the individual practitioner, and reported lab values may need adjustment based on the time period in which they were reported due to changes in equipment calibration and assays.<sup>24</sup> Similar discordance

between definitions was also observed in a comparison of phenotype definitions for diabetes mellitus.<sup>18</sup> Arguably, each phenotype definition may be well suited to a particular purpose, based upon the availability of data, implementation complexity and the need for sensitivity or specificity. To support a LHS operating a comprehensive population management program in CKD, a number of different phenotype definitions may be relevant.

In addition to identifying the full population of patients with any given condition, learning health systems must employ a web of inter-related disease registries and indicators, whose inclusion criteria are defined by computable phenotypes. These registries should identify antecedent and associated risk factors (which may be targets for preventing disease onset or progression), as well as downstream complications. The need for phenotype standardization becomes particularly acute when considering the number of risk factors and complications associated with any given condition.

Portals such as PheKB can support the sharing and thoughtful reuse of phenotype definitions by providing detailed implementation specifications, best practices, and assessments of performance from validation in different settings.<sup>9</sup> Such portals may greatly reduce the cost of implementation. Because the development and validation of computable phenotype definitions is quite resource intensive (often with expert review and/or the ascertainment of an external gold standard), organizations would benefit from reuse of existing definitions whenever possible.



## 4.2 | Standards for clinical decision support

Clinical decision support (CDS) tools use structured rules, developed according to specialty guidelines and subject matter expert input, to improve the quality of care provided to an individual patient, often via prompts presented to the clinician at the point of care in the EHR. CDS tools aim to reduce variation in practice and eliminate divergence from evidence-based care. In one sense, CDS distills the essence of subspecialty expertise to a tractable form that can be used by generalist providers, for example, in primary care.

CDS standards have been proposed to codify the internal activation criteria for a particular decision support tool (ie, for which patients or encounters it should fire), the representation of possible clinical interventions, and the decision logic for recommending one possible intervention over another. The ISA furnishes recommendation on standards for sharable clinical decision support, citing specific HL7, and FHIR resources with supporting information on implementation maturity, cost, and adoption level (relevance in the specific application).<sup>25</sup> This gives insight into the structure of the data and a framework for transmission. However, the identification of patient criteria and triggering events still relies on professional societies to develop consensus dataset and phenotype features.

In the case of CKD, abundant literature indicates that intervention upon CKD risk factors and complications can improve the quality of care and patients' quality of life. Examples of interventions include timely referral to nephrology specialists to prepare for dialysis and/or transplant; management of risk factors such as hypertension, proteinuria, and cardiovascular disease; treatment of anemia and osteodystrophy; and avoidance of medications that are potentially harmful in the setting of CKD. A LHS approach may help to facilitate the identification of the most appropriate standards relevant to these concepts, but it is still essential to identify the governance of these definitions and associate standards among nationally recognized professional societies.

Shiffman et al<sup>26</sup> proposed a methodology for organizing clinical practice guidelines into discrete computable concepts for the purpose of incorporation into CDS tools. This work is relevant to the selection of standards as clinical concepts (whether located in phenotype definitions, clinical guidelines, or patient outcomes), but can often be ambiguous and difficult to define with a succinct set of standards (Table 2).

**TABLE 2** Selected concepts in chronic kidney disease (CKD) decision support and associated standards with assessment of ambiguity

Concept	Standards Domain	Concept Clearly Defined?
Treatment of proteinuria with ACE inhibitor/ARB medication	Clinical terminologies (eg, SNOMED CT, LOINC, RxNorm)	Yes
Avoidance of nephrotoxic medications	Clinical terminologies (eg, RxNorm, LOINC)	Yes
Receipt of CKD education, including preferences for type of renal replacement therapy or palliative care	Clinical terminologies; Patient reported outcomes	No

## 4.3 | Standards for patient education, self-reported outcomes, and care planning

In current practice, EHRs rarely capture discrete data concerning patient preferences, self-reported outcomes, and care planning, reflecting an inherent bias toward applications for billing rather than for providing patient-centric care. As attention shifts toward the patient's experience of care, it will become increasingly important to measure and intervene upon factors that actually matter most to patients, which may diverge from provider-centric measures such as hospitalization and mortality rates.<sup>27,28</sup>

Several validated kidney disease-specific instruments capture patient-reported outcomes, such as the Chronic Kidney Disease Self Efficacy (CKD-SE), Kidney Disease Quality of Life (KDQOL), and Dialysis Symptom Index (DSI) surveys.<sup>29-31</sup> While valid for a particular condition, the administration of disease-specific instruments quickly become burdensome and impractical when solicited for multiple conditions. Many general instruments have developed for recording patient symptoms and purport to be agnostic of underlying health conditions. For example, the PROMIS measures capture patients' symptoms and quality of life indicators using item response theory. Other instruments and vocabularies exist, without a clearly adopted standard. These patient-centric, general instruments offer simplicity and the possibility of reuse among patients suffering from a variety of illnesses, with the caveat that the generic instrument may not have been validated within a narrow population suffering from a specific disease.

Agreement on standards surrounding patient education and care planning is also lacking. For example, patients approaching end-stage renal disease must select from among five treatment options, where the optimal choice depends in large part on patient preference. The options include in-center hemodialysis, home hemodialysis, home peritoneal dialysis, pre-emptive transplantation, and nondialysis conservative management (palliation). The National Kidney Disease Education Program has developed a draft data set for CKD care planning,<sup>32</sup> which includes a LOINC "panel" to capture within the EHR patient preferences regarding renal replacement therapy.<sup>33</sup> To our knowledge, there is no published literature on the usage of this tool in clinical practice. The ongoing PCORI-funded PREPARE NOW study is testing a comprehensive, technology-enabled care planning intervention at Geisinger Health System, and may ultimately produce de facto data standards for this application.<sup>34</sup>

## 4.4 | Standards for measuring quality

Quality measures are a type of standard that seek to measure the quality of medical care in a consistent manner across health care organizations and settings. Quality measures fundamentally involve two components: a denominator, which defines the population to which the measure is applied; and a numerator, which defines the quality criteria to be met.

Mendu and colleagues<sup>17</sup> propose a collection of quality measures for CKD that assess a variety of process and outcome indicators. These measures are derived from professional society guideline recommendations concerning the frequency of monitoring, treatment of complications and comorbidities and preparation for renal replacement therapy among patients progressing to ESRD.

However, quality measures may fail to recognize what matters most to patients.<sup>27,28</sup> Kidney-related quality measures have been criticized for placing undue focus on the processes of care (laboratory monitoring and treatment of comorbidities) and measures of effectiveness (eg, mortality and hospitalization rates), rather than health related quality of life. As in the fable of the man searching at night for his lost car keys under the street lamp because that's the only place he can see, quality measures often reflect data and standards that are convenient to define without ambiguity (see Appendix). Unfortunately, health-related quality of life is not collected consistently or in a standardized way across health organizations, limiting the ability to develop national quality measures based on these outcomes. The Kidney Disease Quality of Life (KQDOL) instrument and others have been used extensively in individuals with ESRD, who in the United States receive care principally from large, national dialysis organizations operating under significant regulatory mandates that include a requirement to collect and incorporate patient-report outcome data into plans of care.<sup>35</sup> In contrast, collection of patient reported outcomes in CKD (ie, prior to progression to dialysis) remains underused and relatively understudied.<sup>36</sup> Until meaningful patient outcomes are routinely captured in EHR systems, quality evaluation using EHR data will likely continue to focus on process measures that are easily captured in transactional systems.

## 5 | DISCUSSION

As a ubiquitous condition defined by objective laboratory criteria and having evidence-based guidelines for treatment, CKD is a model use case for understanding the foundational data standards for population health management in learning health systems. Despite financial incentives that increasingly reward coordinated care, pervasive deficiencies in CKD treatment continue.<sup>37</sup> For any condition, complete, well-specified standards will be needed to define the patient population, to codify the health services delivered, and to evaluate patient-centered outcomes. The application of population health frameworks can leverage these standards to develop scalable informatics tools and care models. In the case of CKD, Mendu et al<sup>17</sup> have created such a framework to advocate for improved care for this population. This existing work provides a basis for identifying standards that can support LHS activities for kidney patients. Our analysis of potential knowledge representation standards in CKD reveals clear gaps in the application and maturity of such standards. Our work highlights the inherent challenges of selecting and implementing standards to undergird a single chronic disease management program. In turn, this may provide generalizable insights regarding the standards needed to support the broader spectrum of LHS activities.

Improved data standards are needed to begin to address these issues pragmatically. First, we need consensus EHR-based phenotype definitions of diseases and conditions. Sources of phenotypes are plentiful, but often contradictory, and almost always contain differing query logic, leading to noncomparable groupings of patients.<sup>9</sup> Second, we need improved, unambiguous data standards to assess quality of care and to develop decision support tools. For example, while standards are available to group medications by class (eg, to track the proportion of patients with proteinuria receiving an angiotensin

converting enzyme [ACE] inhibitor or angiotensin II receptor blocker [ARB]), prescription data are notoriously incomplete when ambulatory EHRs and pharmacy systems are not well integrated. Generic patient reported outcome standards are being developed, such as PROMIS, but may fail to succinctly capture the full burden of symptoms experienced by patients with multisystem illnesses like CKD. Ongoing efforts hold promise for capturing patient preferences and documenting care plans in a structured manner, though have yet to be validated in a real-world setting. To be useful, the design of such standards must be flexible to support reuse across multiple disease applications, and yet sufficiently rigid to avoid issues of ambiguity and duplication present in some of the standards used today.

Our analysis of this case shows that many of the standards needed to set up a disease-focused LHS based upon an established set of treatment and quality measurement recommendations remain underspecified. The ISA provides a strong foundation by naming vocabulary, code set, and terminology standards, but clinical domain groups need to define subsets (ie, value sets) that are relevant to a given condition. The NLM Values Set Authority Center<sup>38</sup> hosts the CMS eCQM and the HL-7 C-CDA values sets, among others. But for a given condition, there are many value sets to choose from, and many have not been validated or assessed for quality and completeness. Further, domain-specific clinical common data elements are lacking for most diseases. The NLM CDE Resource Portal maintains a repository of such elements, but remains incomplete.<sup>39</sup> Ultimately, clinical and patient communities will need to collaborate to identify which standards and subsets are relevant to particular conditions.

These gaps underscore multiple unresolved policy issues around the governance, maintenance, and communication of standards. Who should be responsible for selecting and maintaining standards? How should competing standards be evaluated? What is the optimal mechanism for incorporating technical standards into health care operations? What role should incentives and mandates play? It is difficult to identify in advance a path forward, and a detailed discussion of potential solutions is beyond the scope of this article. In considering the complexities of this single use case of CKD, it is our contention that a centralized, top-down approach to standards governance is not feasible due to the localized and highly contextual nature of health care. We suggest that technical standards organizations and coordinating entities (such as HL7, ONC, and NIH Collaboratory), focus on foundational, "building block" infrastructure (ie, common data elements, evaluation tools, and content libraries). Clinical and patient groups (such as medical specialty societies) could use these standardized tools to collaboratively author and validate higher order archetypes that could be shared and reused by end users for specific applications.

Collaboration and adoption of important standards is needed to alleviate the present issues of scalability of evidence-based medicine and evidence generating medicine. A solution will include an open network to learn from others, an ability to apply existing knowledge and tools, and an effort to engage multiple organizations in the development of a collaborative infrastructure. Stemming from this solution will be adept systems that generate new data, information, and knowledge as the scientific understanding of health care delivery grows.

Continued work and collaboration is required to ensure that a LHS framework can be scalable, moving towards a national LHS network built from many individual localized systems.

Professional societies, organizations, and key stakeholders can collaborate to promote better treatment and rapid learning techniques to increase national knowledge regarding specific conditions and patient centered management approaches.

## 6 | CONCLUSION

Data and knowledge representation standards can support the sharing and reuse of tools that support various LHS activities. The identification and promotion of disease-specific data standards will support the sharing of knowledge and tools that will enable more organizations to undertake LHS activities and thereby support large scale population health management efforts.

## CONFLICT OF INTEREST

The authors have no conflicts of interest to disclose.

## ORCID

Blake Cameron  <http://orcid.org/0000-0002-3063-2880>

## REFERENCES

- Centers for Medicare & Medicaid Services. CMS' Value-Based Programs 2017. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs.html>.
- Bode I, Lange J, Märker M. Caught in organized ambivalence: institutional complexity and its implications in the German hospital sector. *Public Manag Rev*. 2017;19(4):501-517.
- Giussi MV, Baum A, Plazzotta F, Muguerza P, Gonzalez Bernaldo de Quiros F. Change management strategies: transforming a difficult implementation into a successful one. *Stud Health Technol Inform*. 2017;245:813-817.
- Hoog E, Lysholm J, Garvare R, Weinehall L, Nystrom ME. Quality improvement in large healthcare organizations. *J Health Organ Manag*. 2016;30(1):133-153.
- Ngugi BK, Harrington B, Porcher EN, Wamai RG. Data quality shortcomings with the US HIV/AIDS surveillance system. *Health Informatics J*. 2017. 1460458217706183
- Widmer MA, Swanson RC, Zink BJ, Pines JM. Complex systems thinking in emergency medicine: a novel paradigm for a rapidly changing and interconnected health care landscape. *J Eval Clin Pract*. 2017. <https://doi.org/10.1111/jep.12862>
- Health level 7. HL7 EHR behavioral health functional profile, release 1. 2014. [https://www.hl7.org/implement/standards/product\\_brief.cfm?product\\_id=14](https://www.hl7.org/implement/standards/product_brief.cfm?product_id=14)
- Office of the National Coordinator for Health Information Technology. Standards Advisory 2016. <https://www.healthit.gov/standards-advisory>.
- Richesson RL, Smerek MM, Blake CC. A framework to support the sharing and reuse of computable phenotype definitions across health care delivery and clinical research applications. *eGEMs*. 2016;4(3):1232.
- Sittig DF, Wright A, Osheroff JA, et al. Grand challenges in clinical decision support v10. *J Biomed Inform*. 2008;41(2):387-392.
- Institute of Medicine Roundtable on Evidence-Based M. The national academies collection: reports funded by national institutes of health. In: Olsen LA, Aisner D, McGinnis JM, eds. *The Learning Healthcare System: Workshop Summary*. Washington (DC): National Academies Press (US) National Academy of Sciences; 2007.
- Greene SM, Reid RJ, Larson EB. Implementing the learning health system: from concept to action. *Ann Intern Med*. 2012;157(3):207-210.
- U.S. government accountability office. National Institutes of Health: Kidney Disease Research Funding and Priority Setting 2017. <https://www.gao.gov/products/GAO-17-121>.
- Drawz PE, Archdeacon P, McDonald CJ, et al. CKD as a model for improving chronic disease care through electronic health records. *Clin J Am Soc Nephrol: CJASN*. 2015;10(8):1488-1499.
- Stokes JB. Consequences of frequent hemodialysis: comparison to conventional hemodialysis and transplantation. *Trans Am Clin Climatol Assoc*. 2011;122:124-136.
- United States renal data system. 2017 Annual Data Report 2017. <https://www.usrds.org/adr.aspx>.
- Mendu ML, Waikar SS, Rao SK. Kidney disease population health management in the era of accountable care: a conceptual framework for optimizing care across the CKD spectrum. *Am J Kidney Dis*. 2017;70(1):122-131.
- Richesson RL, Rusincovitch SA, Wixted D, et al. A comparison of phenotype definitions for diabetes mellitus. *J Am Med Inform Assoc*. 2013;20(e2):e319-e326.
- Wang V, Maciejewski ML, Hammill BG, et al. Recognition of CKD after the introduction of automated reporting of estimated GFR in the Veterans Health Administration. *Clin J Am Soc Nephrol*. 2014;9(1):29-36.
- Kidney Disease. Improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl*. 2013;3(1):5-14.
- Navaneethan SD, Jolly SE, Schold JD, et al. Development and validation of an electronic health record-based chronic kidney disease registry. *Clin J Am Soc Nephrol: CJASN*. 2011;6(1):40-49.
- Chronic conditions data warehouse. Condition Categories 2018. <https://www.ccwdata.org/web/guest/condition-categories>. Accessed January 21, 2018.
- Cameron CB, Stanifer J, Richesson R. A comparison of electronic health record (EHR) phenotype definitions for chronic kidney disease (CKD) [abstract]. In: American Society of Nephrology Kidney Week; New Orleans, LA. 2017.
- McDonald CJ, Overhage JM. Guidelines you can follow and can trust. An ideal and an example. *JAMA*. 1994;271(11):872-873.
- Office of the National Coordinator for Health Information Technology. Sharable Clinical Decision Support 2017. <https://www.healthit.gov/isa/sharable-clinical-decision-support>.
- Shiffman RN, Michel G, Essaihi A, Thornquist E. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. *J Am Med Inform Assoc*. 2004;11(5):418-426.
- Nissenson AR. Improving outcomes for ESRD patients: shifting the quality paradigm. *Clin J Am Soc Nephrol: CJASN*. 2014;9(2):430-434.
- Urquhart-Secord R, Craig JC, Hemmelgarn B, et al. Patient and caregiver priorities for outcomes in hemodialysis: an international nominal group technique study. *Am J Kidney Dis*. 2016;68(3):444-454.
- Lin C-C, Wu C-C, Anderson RM, et al. The chronic kidney disease self-efficacy (CKD-SE) instrument: development and psychometric evaluation. *Nephrol Dial Transplant*. 2012;27(10):3828-3834.
- RAND corporation. Kidney disease quality of life instrument (KDQOL). 2018. [https://www.rand.org/health/surveys\\_tools/kdqol.html](https://www.rand.org/health/surveys_tools/kdqol.html). Accessed February 15, 2018.
- Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the dialysis symptom index. *J Pain Symptom Manage*. 2004;27(3):226-240.
- National Institute of Diabetes and Digestive and Kidney Diseases. Development of an Electronic CKD Care Plan 2018. <https://www.niddk.nih.gov/health-information/communication-programs/nkdep/working-groups/health-information-technology/development-electronic-ckd-care-plan>. Accessed February 12, 2018.

33. Logical observation identifiers names and codes. Renal replacement therapy goals panel NKDEP 2017. <https://s.details.loinc.org/LOINC/85597-3.html?sections=Comprehensive>.
34. PREPARE NOW. About Us 2015. <http://www.kidneypreparenow.org/about.html>.
35. Flythe JE, Powell JD, Poulton CJ, et al. Patient-reported outcome instruments for physical symptoms among patients receiving maintenance dialysis: a systematic review. *Am J Kidney Dis*. 2015;66(6):1033-1046.
36. Soni RK, Weisbord SD, Unruh ML. Health-related quality of life outcomes in chronic kidney disease. *Curr Opin Nephrol Hypertens*. 2010;19(2):153-159.
37. Jardine MJ, Kasiske B, Adu D, et al. Closing the gap between evidence and practice in chronic kidney disease. *Kidney Int Suppl*. 2017;7(2):114-121.
38. U.S. National Library of Medicine. Value Set Authority Center 2018. <https://vsac.nlm.nih.gov/>. Accessed February 16, 2018.
39. U.S. National Library of Medicine. Common Data Element (CDE) Resource Portal 2018. <https://www.nlm.nih.gov/cde/>. Accessed February 16, 2018.
40. Office of the National Coordinator for Health Information Technology. Representing Patient Tobacco Use (Smoking Status) 2017. <https://www.healthit.gov/isa/representing-patient-tobacco-use-smoking-status>.
41. National Committee for Quality Assurance. NCQA PCMH Standards and Guidelines 2017. <http://store.ncqa.org/index.php/catalog/product/view/id/2776/s/2017-pcmh-standards-and-guidelines-epub/>.
42. Centers for Medicare & Medicaid Services. Stage 2 eligible professional meaningful use core measures, Measure 15 of 17.2015.
43. Office of the National Coordinator for Health Information Technology. Controlling High Blood Pressure 2016. <https://ecqi.healthit.gov/ecqm/measures/cms165v4>.

**How to cite this article:** Cameron B, Douthit B, Richesson R. Data and knowledge standards for learning health: A population management example using chronic kidney disease. *Learn Health Sys*. 2018;2:e10064. <https://doi.org/10.1002/lrh2.10064>

## APPENDIX A

**TABLE A2** Standards Specificity and Coverage of CKD Quality Measures

Proposed CKD Quality Measure from Mendu et al <sup>17</sup>	Specificity and Coverage of Relevant Standards from ISA
Annual eGFR monitoring, reported as percentage of patients with CKD stages 3-5 who had at least 1 eGFR measured in the past 365 days.	Numerator well-specified. LOINC to identify eGFR test result
Annual proteinuria monitoring, reported as percentage of patients with CKD stages 3-5 with at least 1 urine protein measurement in the past 365 days.	Numerator well-specified. LOINC to identify proteinuria test result
Annual assessment of nonsmoking rates, reported as percentage of patients with CKD stages 3-5 with documentation of smoking status in the clinical record indicating never or former smoker in the past 365 days.	Numerator well-specified. SNOMED CT and/or LOINC to assess smoking status per Meaningful Use standard <sup>40</sup>
Completion of nephrology referrals for patients with advanced CKD, reported as percentage of patients with CKD stages 4-5 who have a documented referral, scheduled appointment, or completed visit with a nephrologist in the past 365 days.	Numerator underspecified. Professional guidelines establish varying thresholds for referral, without a clear standard.  Referral communication partially covered under 2017 NCQA Patient-centered Medical Home Care Coordination Standards <sup>41</sup> and Meaningful Use Stage 2 Core Measure 15 standards. <sup>42</sup>  No standard to ensure completed referral or closed loop communication.
Use of ACE inhibitor or ARB therapies for patients with CKD with diabetes mellitus, hypertension, or severely increased albuminuria, reported as percentage of ACE inhibitor/ARB-eligible patients with CKD stages 3-5 with a prescription, refill, or medication verification for ACE inhibitor or ARB in the past 365 days.	Numerator well-specified. LOINC to identify albuminuria RxNorm to identify ACEi/ARB drug class  Denominator underspecified (no clear standard for diabetes mellitus phenotype)
Inappropriate use of unsafe medications in renal patients, reported as percentage of patients with CKD stages 3-5 with at least 1 medication improperly dosed (includes metformin, lithium, Neurontin, atenolol, bisphosphonates, digoxin, nitrofurantoin, Rivaroxaban, trimethoprim-sulfamethoxazole, and levofloxacin) in the past 365 days.	Numerator underspecified  RxNorm to identify unsafe meds (inclusion in of medications in the list is not universally agreed upon and "improper dosing" is ambiguous)
Completion of PROMs, reported as percentage of patients with CKD stages 3-5 with PROMs captured in the EHR in the past 365 days.	Not specified in the ISA  Potential instruments include KDQOL, PROMIS. No established standard.
Blood pressure control, reported as percentage of patients with CKD stages 3-5 with controlled blood pressure ((definitions given)) in the past 365 days.	Numerator underspecified. CMS eCQI 165v4 <sup>43</sup> (multiple competing measures, with variation from current clinical HTN guidelines)

(Continues)



**TABLE A2** (Continued)

Proposed CKD Quality Measure from Mendu et al <sup>17</sup>	Specificity and Coverage of Relevant Standards from ISA
Limit erythrocyte stimulating agent (ESA) overuse, reported as percentage of patients with CKD stages 3-5 for whom an ESA administered in the past 365 days and most recent hemoglobin, 11.5 g/dL measured within 30 days before ESA administration.	Numerator well specified.  LOINC to identify abnormal hemoglobin results RxNorm to identify erythrocyte stimulating agent usage.
Ensure dialysis informed decision making, reported as percentage of patients initiating dialysis therapy whose nephrologists attest within the EHR to completing informed decision-making process in the past 365 days.	Numerator underspecified.  No standard means to capture informed decision making or education. LOINC panel 85597-3 has been proposed.
Ensure timely fistula placement, reported as percentage of patients initiating dialysis therapy whose nephrologists attest within the EHR that the patient has a working fistula in the past 365 days.	Underspecified.  These data are captured on Medicare End-stage Renal Disease Evidence Report Form 2728. No EHR-based standard exists to our knowledge.
All-cause mortality, reported as percentage of patients with CKD stages 3-5 who died in the past 365 days.	Underspecified.  Unreliable electronic capture of out-of-hospital death data in real-world practice.