

7-5-2016

A Framework to Support the Sharing and Re-Use of Computable Phenotype Definitions Across Health Care Delivery and Clinical Research Applications

Rachel Richesson

Duke University School of Nursing, rachel.richesson@dm.duke.edu

Michelle Smerek

Duke Clinical Research Institute, Durham, NC, michelle.smerek@dm.duke.edu

C. Blake Cameron

Duke University School of Medicine, blake.cameron@dm.duke.edu

Follow this and additional works at: <http://repository.edm-forum.org/egems>



Part of the [Health Information Technology Commons](#)

Recommended Citation

Richesson, Rachel; Smerek, Michelle; and Cameron, C. Blake (2016) "A Framework to Support the Sharing and Re-Use of Computable Phenotype Definitions Across Health Care Delivery and Clinical Research Applications," *eGEMs (Generating Evidence & Methods to improve patient outcomes)*: Vol. 4: Iss. 3, Article 2.

DOI: <https://doi.org/10.13063/2327-9214.1232>

Available at: <http://repository.edm-forum.org/egems/vol4/iss3/2>

This Learning Health System Model/Framework is brought to you for free and open access by the the Publish at EDM Forum Community. It has been peer-reviewed and accepted for publication in eGEMs (Generating Evidence & Methods to improve patient outcomes).

The Electronic Data Methods (EDM) Forum is supported by the Agency for Healthcare Research and Quality (AHRQ), Grant 1U18HS022789-01. eGEMs publications do not reflect the official views of AHRQ or the United States Department of Health and Human Services.

A Framework to Support the Sharing and Re-Use of Computable Phenotype Definitions Across Health Care Delivery and Clinical Research Applications

Abstract

Introduction: The ability to reproducibly identify clinically equivalent patient populations is critical to the vision of learning health care systems that implement and evaluate evidence-based treatments. The use of common or semantically equivalent phenotype definitions across research and health care use cases will support this aim. Currently, there is no single consolidated repository for computable phenotype definitions, making it difficult to find all definitions that already exist, and also hindering the sharing of definitions between user groups.

Method: Drawing from our experience in an academic medical center that supports a number of multisite research projects and quality improvement studies, we articulate a framework that will support the sharing of phenotype definitions across research and health care use cases, and highlight gaps and areas that need attention and collaborative solutions.

Framework: An infrastructure for re-using computable phenotype definitions and sharing experience across health care delivery and clinical research applications includes: access to a collection of existing phenotype definitions, information to evaluate their appropriateness for particular applications, a knowledge base of implementation guidance, supporting tools that are user-friendly and intuitive, and a willingness to use them.

Next Steps: We encourage prospective researchers and health administrators to re-use existing EHR-based condition definitions where appropriate and share their results with others to support a national culture of learning health care. There are a number of federally funded resources to support these activities, and research sponsors should encourage their use.

Acknowledgements

The SEDI project and ancillary study described was supported by Cooperative Agreement Number 1C1CMS331018-01-00 from the Department of Health and Human Services, Centers for Medicare & Medicaid Services. This publication was also made possible by the Patient Centered Outcomes Research Institute (PCORI) and the National Institutes of Health (NIH) Common Fund, through a cooperative agreement (U54 AT007748) from the Office of Strategic Coordination within the Office of the NIH Director and the Duke CTSA (UL1TR001117). Dr. Cameron is supported by grant 5 T32 DK007731 (Duke Training Grant in Nephrology) The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the Patient Centered Outcomes Research or the U.S. Department of Health and Human Services or any of its agencies. We are grateful to Shelley Rusincovitch of the Duke Translational Medicine Institute for her ideas in the formulation of this commentary and to the members of the PSQ Core of the NIH Collaboratory for their ideas and discussion related to achieving valid and scalable phenotype definitions for a number of conditions.

Keywords

Computable Phenotypes; Electronic Health Records; Data Standards; Learning Health Care Systems

Disciplines

Health Information Technology

Creative Commons License

This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).



A Framework to Support the Sharing and Reuse of Computable Phenotype Definitions Across Health Care Delivery and Clinical Research Applications

Rachel L. Richesson, PhDⁱ; Michelle M. Smerekⁱⁱ; C. Blake Cameron, MDⁱⁱⁱ

ABSTRACT

Introduction: The ability to reproducibly identify clinically equivalent patient populations is critical to the vision of learning health care systems that implement and evaluate evidence-based treatments. The use of common or semantically equivalent phenotype definitions across research and health care use cases will support this aim. Currently, there is no single consolidated repository for computable phenotype definitions, making it difficult to find all definitions that already exist, and also hindering the sharing of definitions between user groups.

Method: Drawing from our experience in an academic medical center that supports a number of multisite research projects and quality improvement studies, we articulate a framework that will support the sharing of phenotype definitions across research and health care use cases, and highlight gaps and areas that need attention and collaborative solutions.

Framework: An infrastructure for re-using computable phenotype definitions and sharing experience across health care delivery and clinical research applications includes: access to a collection of existing phenotype definitions, information to evaluate their appropriateness for particular applications, a knowledge base of implementation guidance, supporting tools that are user-friendly and intuitive, and a willingness to use them.

Next Steps: We encourage prospective researchers and health administrators to re-use existing EHR-based condition definitions where appropriate and share their results with others to support a national culture of learning health care. There are a number of federally funded resources to support these activities, and research sponsors should encourage their use.

ⁱDuke University School of Nursing, ⁱⁱDuke Clinical Research Institute, ⁱⁱⁱDuke University School of Medicine

Introduction

Computable phenotypes, or electronic health record (EHR)-based condition definitions, enable the identification of cohorts of patients with certain diseases or clinical profiles for disease management registries, quality improvement programs, evaluation studies, and interventional research. Regardless of the application, cohort identification requires queries of clinical data stores that are both valid and reproducible. Currently, there is no single consolidated repository for computable phenotypes, making it difficult to find all definitions that already exist, and also hindering the sharing of definitions between user groups. Health services researchers and quality assessment groups—i.e., the National Quality Forum (NQF), National Committee for Quality Assurance (NCQA), the Centers for Medicare & Medicaid Services (CMS), and the Agency for Healthcare Research and Quality (AHRQ)—provide computable phenotypes on a number of websites.¹⁻⁵ In addition, researchers and registry developers create definitions utilizing different design and evaluation methods. Because the definitional logic is often underspecified or unreported in scientific journals, it is not clear if the findings reported in published research or quality improvement are comparable or relevant to clinical populations, hindering the application of evidence-based medical and nursing care.

We believe that a minimal set of well-constructed and explicit EHR-based phenotype definitions will create efficiencies for health care organizations that must increasingly support growing numbers of data requests related to comparative effectiveness research (CER), quality improvement, and chronic disease management. We further believe that such a set will facilitate synergies between research and care delivery, enabling “learning health care” practices⁶ and subsequently improving patient outcomes. A large-scale and multipurpose approach

to sharing phenotype definitions will support the reuse of well-constructed and validated computable phenotypes, and will subsequently reduce the variation in definitions across conditions. Drawing from our experience from an academic medical center supporting a number of multisite research projects, we articulate a framework that will support the sharing of phenotype definitions across research and health care use cases, and highlight gaps or areas that need attention and collaborative solutions.

Background and Context

A “computable phenotype” is a definition of a condition, disease, or characteristic or clinical event that is based solely on data that can be processed by a computer.⁷ Computable phenotype definitions provide the specifications to identify populations of patients with conditions of interest, and can be combined with other criteria, such as age or other demographic information, to develop cohort populations for a variety of purposes.

Quality monitoring organizations (such as NQF, NCQA, and AHRQ) create computable phenotype definitions for the development and monitoring of health care quality measures. A number of research networks have developed phenotype definitions to enable the use of EHR data for observational research (including comparative effectiveness studies) and interventional trials.⁸⁻¹¹ Various multisite studies^{12,13} use these definitions to develop registries for drug safety surveillance¹⁴ or chronic disease management.¹⁵ There are numerous and distinct use cases for computable phenotypes for health care delivery (e.g., personalized medicine, guideline-based care, chronic disease management, and quality measurement) and biomedical research (genomic, observational, CER, health services research, and interventional trials.) Each use case represents different scientific disciplines whose phenotype development efforts have heretofore



been undertaken in isolation, without the benefit of cooperation. Further, there are no standards of practice that encourage reuse of existing definitions or the use of common definitions for health care delivery and research uses.

The lack of coordination of phenotype definitions among researchers, clinicians, and administrators has led to the unintentional proliferation of numerous definitions for many conditions and clinical profiles. Because each definition applies different logic (e.g., various combinations of diagnosis or procedure codes, medications, or laboratory tests) for querying EHR data, the resulting cohorts are often not directly comparable. It is unknown how much semantic variation in definitions actually exists, because this information is often underspecified in research publications. A recent report on national trends in diabetes specifically lists several related conditions (including hypoglycemia, neuropathy, chronic kidney disease, peripheral vascular disease, cognitive decline, cancers, and even differentiating type 1 from type 2 diabetes) whose prevalence could not be reported due to inconsistent EHR documentation and definitions across the United States.¹⁶ The consequent likelihood that research, patient care, and quality measurement communities are using different phenotype definitions for the same condition is more concerning. The COPD Outcomes-based Network for Clinical Effectiveness & Research Translation (CONCERT) assessed 980 patients sampled from various EHR systems using a clinical phenotype definition for chronic obstructive pulmonary disease (COPD), and found that just over half of those met the criteria for the well-accepted research definition for the condition.¹⁵ Further, they found that the patient populations retrieved by the clinical and research definitions for COPD had significantly different comorbidities and risk factors.¹³ This implies that disease management registries and

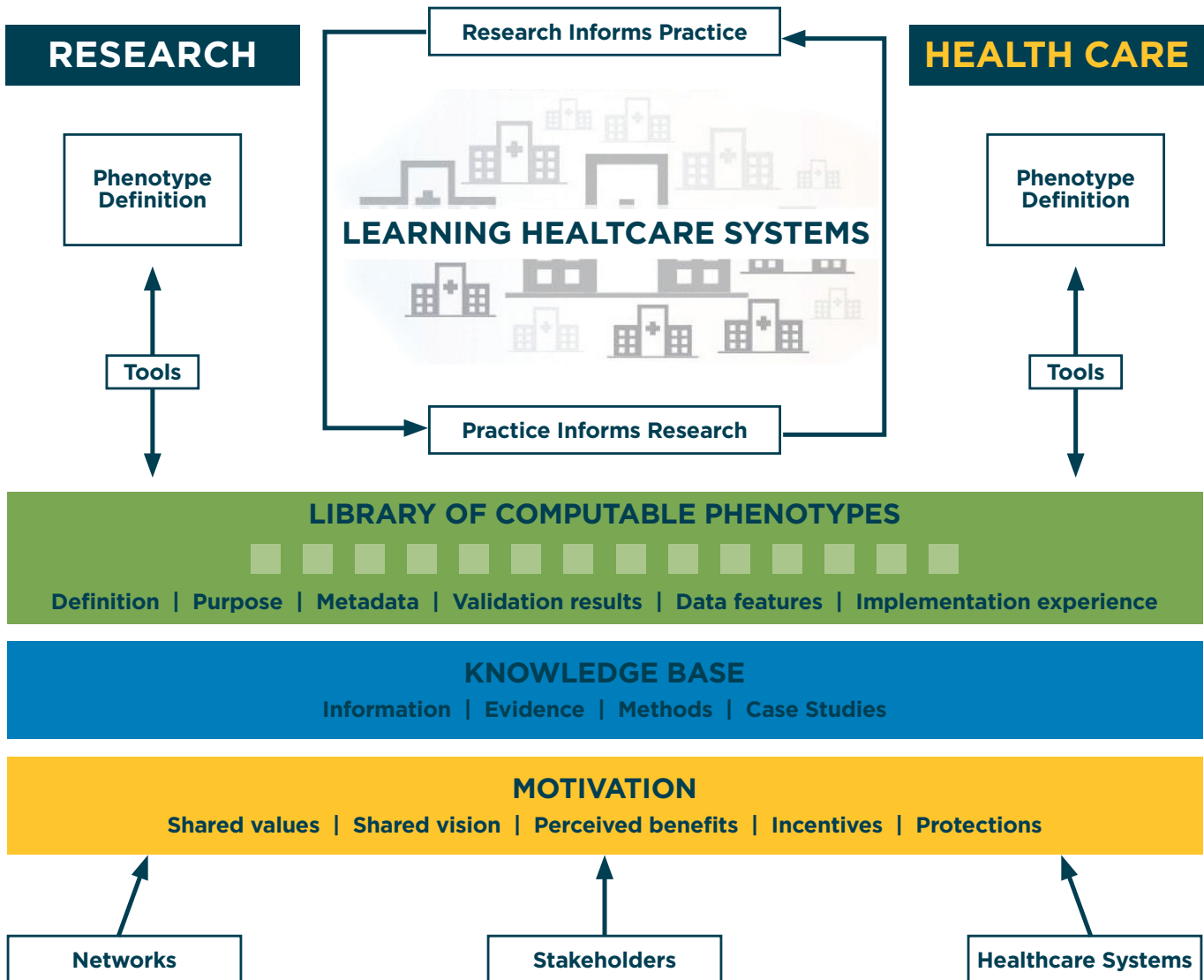
quality improvement programs might be identifying populations that are different from those used in the development of the evidence upon which those supporting treatment strategies and interventions are based.

The “research informs practice informs research” cycle that is the essence of learning health care systems entails that the clinical features used to define research and patient populations be well understood and comparable. Hence semantically equivalent phenotype definitions must be used to identify clinically equivalent populations. We believe that creating a centralized collection of explicitly defined computable phenotypes, with an accompanying knowledge base of development and validation documentation, is the first step toward consolidating effort and harmonizing definitions. Information, resources, and tools that facilitate the reuse of existing phenotypes will reduce the variation in phenotype definitions across all use cases, facilitate conversations between health care and research communities about how to compare definitions for different use cases, and ultimately lead to harmonization of definitions that will simplify and support the identification of clinically equivalent populations for research and health care purposes.

Framework Components

The reuse of phenotype definitions can be facilitated by their explicit representation and tools to support their evaluation and implementation in new applications. We propose that the deliberate and informed reuse of existing definitions will require four components: (1) searchable libraries of explicitly defined phenotype definitions; (2) supporting knowledge bases with information and methods; (3) tools to identify, evaluate, and implement existing phenotype definitions; and (4) motivated users and stakeholders to use them (Fig 1).

Figure 1. Overview of Framework to Support the Reuse of Phenotype Definitions in Learning Health Care Systems





Searchable Libraries of Phenotype Definitions

The sharing of information about computable phenotype definitions will allow implementers to reuse appropriate existing definitions rather than creating their own. This requires access to an ample set of phenotype definitions, along with information that enables them to be evaluated and easily implemented. The ideal library should be indexed so that users can search by a number of different features including the clinical condition; the data elements; logic; the intended use case; limitations; and orientation toward precision, sensitivity or specificity.

Mo and colleagues call for a formal computable representation of phenotype definitions that will enable scalability of the definitions by allowing them to be applied to different data systems.¹⁷ Their desiderata includes the following: human-readable and computable forms, structured rules, formalisms for temporal relations, representations for text searching and natural language processing, and interfaces for external software algorithms. They endorse the use of standardized terminologies, ontologies, and also the reuse of value sets.

Additional information can be included in the library or underlying knowledge base to support users' *semantic understanding* of the phenotype definition, and to enable selection of the appropriate definition to identify patient cohorts with the intended clinical features. Therefore, the definitions in a phenotype library should include metadata or supporting information about a definition, its intended use, the clinical rationale or research justification for the definition, and data about clinical and scientific validation in various health care settings. As an example, actual blood pressure measurements, even when they are available for long periods, did not contribute significantly to predictive models for

hypertension control.¹⁸ Without clear supporting documentation, clinical subject matter experts may reject, as lacking face validity, well-validated phenotype definitions that do not match their expectations or intuition. Clinical practice and disease definitions change over time. Therefore, phenotype definitions in a phenotype library should reference the underlying clinical definitions or guidelines upon which they are based, in order to better identify legacy definitions that are out of date. In addition, phenotype definitions should conform to existing required and emerging terminologies and standards—e.g., SNOMED CT, LOINC, RxNorm, LOINC, NDF-RT—for representing clinical data, as endorsed by the Office of the National Coordinator.¹⁹ Adherence to standards allows for a modular design that reduces development and implementation costs, particularly at scale where multiple use cases for that standard may exist concurrently.

Because phenotype definitions might perform differently when implemented in different patient populations and EHR systems, information about the performance of phenotypes in specific organizations should be collected from implementers and shared with future users. Implementation information is necessary to understand how standard definitions perform across diverse populations, heterogeneous organizations and EHRs systems. Specifically, information about the underlying population and quality (i.e., completeness, accuracy, consistency) of data that were used to validate the definitions have important implications for interpreting the validation results. For example, if a test population had 50 percent missing data in one of the defining variables for the phenotype, the provision of this information provides important contextual information about the definition's performance. Similarly, the testing of phenotype definitions in populations with high versus low prevalence of disease will yield different results. Recommendations for data quality

assessment reporting in pragmatic trials²⁰ and observational research^{21,22} can provide insight into which data quality dimensions (e.g., completeness, accuracy) might be most useful to evaluate the phenotype definition.

To maximize the socialization and collaboration around shared phenotypes, the ideal phenotype library should support communication between phenotype developers and implementers. The Centers for Medicare & Medicaid Services (CMS) employs a standardized approach for enabling users to post questions and share comments, and for maintaining quality measure definitions across multiple programs.²³ Such a framework could be adapted for use with computable phenotype libraries. During early development, draft phenotype specifications could be posted in a library for public review to evaluate feasibility and refine use cases. During the validation phase, the testing methodology could be opened to public comment. Once validated, the library could facilitate communications of best practices and feedback. This would allow implementers to share information about their experiences implementing phenotype definitions in their local systems, and allow others to ask questions to inform the many practical decisions that are made when implementing abstract logic in local data systems. A collaborative or interactive component would also allow users to relate their experience implementing definitions in different vendor systems and in different patient populations. Over time, the library could collect data on usage and impact, and aggregate published literature based on each phenotype. A record of projects that have used or endorsed different phenotype definitions can enhance understanding of phenotype intent and performance, and can assist potential implementers in the selection of appropriate phenotypes.

Because phenotype definitions are dynamic, the library should reference a phenotype life cycle,

the current phenotype life cycle stage, and should include the status (e.g., in development, draft, final), as well as tracking the version number or date of last revision. Phenotypes could be marked as retired or archived in cases where clinical practice changes or the underlying clinical definitions or data standards become out of date.

The Phenotype Knowledge Base (PheKB)²⁴ is a large and well-indexed portal for hosting computable phenotypes, though enhancements are needed to accommodate the above information requirements. PheKB includes human-readable definitions and machine-readable code in some cases, but the code is not fully executable across heterogeneous EHR systems. The PheKB does have an interface for reporting contextual data and performance metrics of phenotype definitions, but a useful and usable display of these data is not yet standardized. Also, it is not known how widely PheKB is used outside of the Electronic Medical Records and Genomics (eMERGE) Network or Pharmacogenetic Research Network, whose goals are to implement decision support around clinically actionable genetic variants for clinical conditions. While several National Institutes of Health (NIH) Collaboratory and National Patient-Centered Clinical Research Network (PCORnet) investigators have added their phenotype definitions to PheKB, an increased uptake of PheKB by other research and clinical groups will require targeted marketing. Broader usage of PheKB might drive enhancements to the PheKB resource, but also will likely increase the number of user requirements. Several authoring tools exist, including the PheMA project,²⁵ which provides generalizable computable representations and automated mapping tools.²⁶ Other research networks, such as Observational Health Data Sciences and Informatics (OHDSI)²⁷ and PCORnet,¹¹ include dedicated phenotype working groups and internal inventories of phenotype definitions.



Knowledge Base of Information and Methods

Researchers and health care organizations need information about how to develop, evaluate, and implement phenotype definitions. The particular use case influences the nature of the phenotype definition and system requirements. For example, in quality measurement, the purpose of the phenotype definition is to identify “bread and butter” instances of a particular condition. Patients whose disease status is negative or indeterminate are excluded. By contrast, genomic research usually aims to reliably identify both cases and controls (negative cases). The phenotype definition must identify with reasonable certainty not only patients who have the condition (i.e., have adequate sensitivity), but also patients who clearly do not have the condition (i.e., high specificity). Definitions used in disease management registries or population health promotion activities have needs for higher sensitivity at the cost of specificity, whereas CER requires higher specificity and precision. Guidance from different health care and research communities can inform users about important features and performance thresholds for phenotype definitions for different use cases.

Information to clarify data dependencies and implementation requirements is needed to facilitate the sharing of phenotypes across groups. For example, some definitions include natural language processing (NLP) components that might not be feasible for some target systems. The information in the knowledge base can include methods and case studies from projects that have implemented the definitions in multiple organizations; their customizations and lessons learned can inform future users. Evidence-based practice guidelines that include justification for a definition’s logic as well as the definition of “gold standard” for validation of EHR-based phenotype definitions should also be available.

Although PheKB does include some “knowledge” in the form of phenotype development methods and validation protocols, it is limited and not tailored for different phenotype users. Information for a broader range of use cases is needed. *Rethinking Clinical Trials: The Living Textbook of Pragmatic Clinical Trials*⁹ provides a model for disseminating information in the form of “lessons learned” and case studies, rather than as empirical research. Many other research-network websites and collaborative networks perform this function, but a central portal to the knowledge from various networks would support potential implementers from multiple domains.

Tools

Formal representations of computable phenotypes, mappings to reference coding systems and (common) information models, and executable code can support the implementation of definitions in different populations. Mo’s desiderata highlights recommendations for clinical data representation to support phenotyping.¹⁷ This specifically calls for the structure of clinical data into queryable forms and the use of a common data model to support customization for the variability and availability of EHR data among sites. Since there currently are a number of (different) common data models used in research networks,²⁸⁻³⁰ there is a need for tools and platforms to implement a given phenotype definition in different contexts. Knowledge, authoring tools, and vocabulary mapping tools to support these activities can also be centrally available through a shared knowledge base³¹ or links to a code sharing base like GitHub. Similarly, the implementation of these definitions require terminology mappings (e.g., from drug class names in NDF-RT and medication sets from RxNorm to product codes in (NDC)).³² Terminology integration resources, such as RxNorm, the Unified Medical Language System (UMLS) and UMLS Terminology Services (UTS) tools, can benefit

phenotype use cases in many networks. To be more broadly used, these tools should be centrally available with supporting instructions for people from many different domains and levels of technical expertise.

Specifically, tools are needed for the following uses: (1) *searching* for phenotype definitions that are endorsed or mandated; (2) *browsing* existing phenotypes to find ones that are potentially relevant that can be reused; (3) *the display of relevant information* to help potential implementers understand existing definitions and their strengths and limitations for particular uses; (4) the *implementation* of those definitions in local EHR systems with, e.g., executable code tailored to

common data models, or mappings between coding systems; (5) developing new phenotype definitions if needed; and (6) *reporting implementation results, along with characteristics of test data sets*) for others to view (Table 1).

We see gaps and unmet needs in all areas except for phenotype development. At least two scalable authoring tools exist—PheMA²⁶ with its execution support³¹ and OHDSI’s CALYPSO (Criteria Assessment Logic for Your Population Study in Observational data).^{33,34} Xu et al. provide a detailed inventory of other search and authoring tools.²⁵ In addition to guided phenotype authoring tools based on the underlying model of the phenotype library, other tools theoretically could support an “import

Table 1. Types of Tools and Functionality Required to Support the Sharing and Reuse of Computable Phenotype Definitions Across Health Care Delivery and Clinical Research Applications

| FUNCTION | PURPOSE | EXAMPLE OR POTENTIAL TOOL |
|---|---|--|
| Search for phenotype definitions. | Identify validated or endorsed phenotype definitions. | PheKB |
| Browse for phenotype definitions. | Assess landscape. | PheKB |
| Display pertinent context information. | Aid potential implementers in assessing a definitions fit for their use case. | needed* |
| Provide executable code in different formats (SQL, SAS, R, etc.) and crosswalks for mapping between different coding systems. | Implement phenotype definitions in heterogeneous systems. | PheKB, |
| GitHub | | |
| Develop new phenotype definitions. | Create new definitions when existing ones aren’t a good fit. | PheMA ²⁶ CALYPSO ³³ |
| Display implementation results with characteristics of the data in which phenotypes were implemented. | Provide additional information users need to consider when determining whether a definition is a good fit for their use case. | needed* |

Note: *This represents a gap where tooling is needed. We are not aware of existing tools that support this function.



and transformation” process that could take existing definitions developed locally (with local tools) and store them in the central repository for other to access and use.

The learning health system cannot exist on phenotypes alone. Any phenotype library would need to provide a service-based API that other computable clinical “services” might be able to access in a standardized way, e.g., electronic clinical-trial management tools that might access existing phenotype definitions in order to define the inclusion and exclusion criteria for a research trial. A number of functional components, e.g., standard models and vocabulary services, would in turn be needed to fully support the reuse of phenotype definitions on a grand scale.

Motivated Users and Stakeholders

The sharing of definitions and experience will require deliberate action on the part of potential phenotype developers and implementers, and useful and intuitive tools can support this behavior. Aligning existing computable phenotypes with users’ needs will likely positively influence their uptake, as will engaging all stakeholders in the design and development of phenotype resources and tools described in this framework. Additionally, a number of approaches can be used to motivate individuals to search for existing definitions and to share the outcomes of computable phenotype implementations. Possible approaches include creating incentives, increasing perceived benefit, establishing new social norms, or regulating with policies or regulations.

Perceived Benefits and Value

Collaboration is fostered when the collaborators expect or perceive a beneficial outcome. The more beneficial or significant the outcome, the higher the participation and commitment level among

collaborators will be. Wilcox et al. assert that the costs for sustaining research infrastructure can be covered if value can be created.³⁵ Thus, clear demonstrations of reduced workload, reduced costs, or faster development resulting from the reuse of phenotype definitions might motivate potential users.

Incentives

Tangible incentives can be created through policy or legislation. Examples include quality reporting incentives (e.g., the CMS Physician Quality Reporting System and the financial rewards of the Meaningful Use program), and punitive consequences for noncompliance with Food and Drug Administration (FDA) reporting specifications. Although these types of incentives might be effective, they are time-consuming and expensive to achieve. Alternative incentives might derive from some sort of peer pressure from the scientific community to report phenotype definitions as part of the research protocol or study results reporting in publications, or rewards for such behavior from research sponsors or in academic promotion rubrics.

Shared Values and Principles

A set of agreed upon assumptions and principles for research networks, sponsors, and health care regulators to adopt is the first step in addressing the complex challenges to reusing phenotype definitions. These should include a stated commitment to reproducible science and the standardized reporting of phenotype definitions, use case, and validation results. Additional principles could include an expectation that users of computable phenotypes will search for and consider existing definitions before creating their own. For conditions where a phenotype definition already exists, researchers should carefully consider whether the benefit of developing new definitions tailored to their specific use cases outweighs the

Box 1. User Scenarios that Illustrate Benefit of Shared Phenotype Definitions

Scenario 1.

An intervention specialist working for the Southeastern Diabetes Initiative (SEDI) wants to identify patients with type 2 diabetes across a number of health care providers in order to develop treatment programs and community interventions that will improve diabetes care. The specialist needs operational definitions for type 2 diabetes, as well as a number of associated conditions such as hypertension and chronic kidney disease. She goes to a central phenotype library and finds definitions for each condition that are appropriate for broad population screening and that can be implemented in all the SEDI sites, including one with no capacity for accessing clinical notes. She shares a link for each selected phenotype definition, plus implementation guidance and appropriate code, with the data specialists at each SEDI site. Each site implements the definition and reports their results to the phenotype library. One SEDI site had problems with the code and reported this experience as well. The original developer of the phenotype contacted the SEDI site with a suggestion. This suggestion was helpful and was therefore added to the knowledge base for other SEDI sites to access and review. Later, the study was published in a journal and referenced the link to the computable phenotype logic and supporting implementation tools. Using these definitions and tools, a new group of researchers replicated the intervention in an urban population on the West Coast and published their findings.

This scenario was enabled by the following:

1. Searchable libraries of explicitly defined phenotype definitions;
2. Supporting knowledge bases with information and methods;
3. Supporting tools; and
4. Users and stakeholders motivated to consider reusing existing definitions; benefits from reuse and shared phenotype definitions were realized by the users.

Scenario 2.

A clinician reviews the literature and finds a study of a new medical intervention for uncontrolled hypertension. She wants to implement it on a similar population in her clinic. The published article includes a narrative discussion of the inclusion and exclusion criteria (e.g., includes diagnosis of hypertension and excludes chronic kidney disease) with hyperlinks to a public phenotype library that hosts the computable phenotype specifications for the intervention population. The clinician points her data analyst to the phenotype specifications and requests a data warehouse query to estimate the number of patients that might be eligible for the planned intervention. After obtaining the required institutional approvals, she implements the intervention and conducts a formal quality improvement study. She publishes that study and references a public link to the phenotype library and knowledge base for the specific computable phenotype-definition logic and supporting implementation tools. Future implementers access the library for implementation details, rather than contacting this clinical investigator, allowing her more time to research and plan new chronic disease management interventions.

This scenario was enabled by the following:

1. Searchable libraries of explicitly defined phenotype definitions;
2. Supporting knowledge bases with information and methods;
3. Supporting tools; and
4. Users and stakeholders motivated to consider reusing existing definitions; benefits from reuse and shared phenotype definitions were realized by the users.



losses incurred by sacrificing interoperability. Other principles might include that phenotype definitions should be placed in the public domain—regardless of whether they derived from federally funded research, national quality reporting incentive programs, or private ventures. Because phenotype definitions are developed for different purposes, populations, and settings, it is not feasible to define a set of definitions for all research needs. However, the explicit documentation and sharing of phenotype definitions and supporting evidence will enable researchers to evaluate and select the best available definitions for their populations and research needs. While there may be potential research integrity risks associated with using data or methods without full understanding of their limitations, a repository with information about the intent, maturity, and limitations of particular phenotype definitions can inform and empower potential users to use them appropriately and at their own prudence.

Vision of Shared Phenotype Definitions

The need for a shared or common vision has been identified as important success factors in collaborative projects. We provide a vision in the form of two scenarios that might motivate pan-network or cross-use case sharing of phenotype definitions (Box 1).

Communication, Marketing, and Engagement

Communication and marketing of a set of principles and vision might enhance the engagement, participation, and support of stakeholders from multiple organizations and domains. Communication campaigns that inform potential users about the availability of existing computable phenotypes and increase their perception that reusing existing definitions will save them work, or produce a better definition (that has been previously tested) than they can do alone. Professional societies and medical advocacy groups may choose to endorse and

curate authoritative phenotypes as a complement to guideline development activities. Further, models of sharing behavior could be manufactured and made visible, such as online exchanges between investigators that describe challenges or observations in implementing particular definitions in certain settings.

Protection from Risks

Inherent in understanding the motivation for sharing is to understand what fears or hesitations research investigators or project implementers might have. Anecdotally, the risks to sharing are concerns about publication, copyright, or inappropriate use. Phenotype developers might not feel their definition is of broad interest, thinking it too institution- or protocol specific to be of interest to other users, or they may have concern that it is not ready. These factors need to be researched and understood in order to create stronger alternative incentives or beliefs that will motivate developers of phenotypes to share their definitions.

Discussion

Computable phenotype definitions that are developed and represented in an explicit and standardized manner are necessary to ensure the consistency of clinical populations sampled for different purposes. The use of semantically equivalent phenotype definitions can enable the comparison of results across studies, and ensure that all patients can be reliably identified and offered evidence-based treatment options and opportunities for research. We do not suggest that a single definition per condition is feasible, nor that one definition per use case will necessarily be sufficient. However, we do suggest that some minimum set of definitions per condition can be identified to address the majority of use cases. It will be important to have resources and communication in place to ensure that the definitions are as accurate and scalable as

possible, and that users can identify the definitions that are the best fit for their intended purposes.

Within research networks, member investigators have a vested interest in maintaining the health of the network, and therefore are well incentivized to support policies, communication channels, and tools that enable and encourage the sharing and reuse of phenotype definitions within the network. Sharing phenotype definitions *across* networks or domains (e.g., from research to health care quality improvement) will be more challenging to motivate, as it involves multiple organizations and complex systems whose incentive structures may differ. Evidence-based methods that support the collaboration of diverse stakeholders to solve challenging problems in complex systems should be applied to support the sharing and standardization of computable phenotypes between health care and research. The lack of supporting theories and methods for complex cross-boundary collaborations³⁶ illustrates a gap in learning health sciences that should be addressed.³⁷

The learning health care paradigm will demand continuous development and refinement of new phenotypes to identify conditions of interest and to reflect changes in health care practice and EHR systems. Clinicians, health care administrators, investigators, and patients benefit from the use of explicitly defined and validated definitions for sampling, potential research participant identification, and broader analyses using data from EHRs. Collaboration around the development of computable phenotypes for emerging diseases, especially where consensus in professional societies is slow to emerge (e.g., the early years of HIV/AIDS) or varies over time, e.g., the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5)'s new classification of the autism spectrum—which is not concordant with prior definitions, might expedite their investigation and

build consensus in professional society guidelines in a rapid learning environment. Similarly, standardized processes to update and periodically revalidate definitions—as knowledge of disease increases, and as coded terminologies, EHRs and patterns of health care delivery mature—will be required.

The creation of a culture for sharing, reusing, and harmonizing phenotype definitions will require changes in thinking and behavior that can be enhanced by the following call to action for researchers and clinicians: (1) champion cultural changes and resource allocations that will enable the reuse of computable phenotype definitions where appropriate; (2) survey the landscape for existing and previously validated definitions that will meet the particular need before creating a new definition, and (3) provide phenotype definition logic and implementation performance or validation results, so that others can benefit from this knowledge.

Box 2. Call to Action for Researchers and Clinicians to Facilitate Learning Health Care Systems

1. Champion cultural changes and resource allocations that will enable the reuse of computable phenotype definitions where appropriate.
2. Survey the landscape for existing and previously validated definitions that will meet the particular need before creating a new definition.
3. Provide phenotype definition logic and implementation performance or validation results, so that others can benefit from this knowledge.

The vision of shared phenotype definitions between research and health care activities will ultimately require governance structures to control curation of phenotype knowledge, raising a number of questions that will need to be addressed: Who should be the guardians of such knowledge—a centrally controlled federal agency or commercial entity, or both? What are the types of criteria that would be used to



accept a phenotype definition into the repository? Specifically, what gold standard evidence-based practice guideline sources are deemed of sufficient quality to be acceptable as a basis for phenotype definition?

The perceived benefits of shared phenotypes might drive funding or advocacy for developing and enhancing resources to support the sharing and reuse of computable phenotype definitions across health care delivery and clinical research applications, but measurable results or return on investment effort will be necessary to maintain them and motivate widespread use in learning health systems. Financial models for phenotype contributors and users will need to be explored. Ultimately, the vision of shared phenotype definitions will only transpire if the libraries, knowledge bases, tools, and processes are usable and useful for users, and if the sharing of definitions creates efficiency for research and health care teams, as well as a synergy between them that benefits patients, payors, and other stakeholders.

Conclusions and Call to Action

The implementation of learning health care systems is gaining momentum, and the ability to reproducibly identify clinically equivalent patient populations is critical to implementing and evaluating evidence-based treatments in health care systems. The use of common or semantically equivalent phenotype definitions across research and health care use cases can support this aim. A national infrastructure for reusing phenotype definitions and sharing experience across health care delivery and clinical research applications will reduce duplicate efforts and increase efficiencies. Both research and provider communities need access to a collection of existing definitions, information to evaluate their appropriateness for particular applications,

a knowledge base of implementation guidance, supporting tools that are user-friendly and intuitive, and a willingness to use them. We encourage prospective researchers and health administrators to reuse existing EHR-based condition definitions where appropriate and to share their results with others to support a national culture of learning health care. A number of federally funded resources support these activities, and research sponsors should encourage their use.

Acknowledgments

The SEDI project and ancillary study described was supported by Cooperative Agreement Number 1C1CMS331018-01-00 from the Department of Health and Human Services, Centers for Medicare & Medicaid Services. This publication was also made possible by the Patient Centered Outcomes Research Institute (PCORI) and the National Institutes of Health (NIH) Common Fund, through a cooperative agreement (U54 AT007748) from the Office of Strategic Coordination within the Office of the NIH Director and the Duke CTSA (UL1TR001117). Dr. Cameron is supported by grant 5 T32 DK007731 (Duke Training Grant in Nephrology)

The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the Patient Centered Outcomes Research or the U.S. Department of Health and Human Services or any of its agencies.

We are grateful to Shelley Rusincovitch of the Duke Translational Medicine Institute for her ideas in the formulation of this commentary and to the members of the PSQ Core of the NIH Collaboratory for their ideas and discussion related to achieving valid and scalable phenotype definitions for a number of conditions.

References

1. NQF. Measure Developer Guide to Submitting Measures to NQF October 2013 2013. Available from: http://www.qualityforum.org/Measuring_Performance/Endorsed_Performance_Measures_Maintenance.aspx.
2. CMS. Chronic condition data warehouse: Centers for Medicare and Medicaid Services 2013 [cited 2013 May 27]. Available from: <http://www.ccwdata.org/index.htm>.
3. CMS. eCQM Library: Centers for Medicare & Medicaid Services; 2015. Available from: https://www.cms.gov/regulations-and-guidance/legislation/ehrincentiveprograms/ecqm_library.html.
4. NCQA. Healthcare Effectiveness Data and Information Set (HEDIS) Measures: National Committee for Quality Assurance; 2015 [cited 2015 September 28]. Available from: <http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures.aspx>
5. CCS AH. Healthcare Cost and Utilization Project (HCUP). Clinical Classifications Software (CCS) for ICD-9-CM. Rockville, MD: Agency for Healthcare Research and Quality 2013. Available from: <http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>.
6. Greene SM, Reid RJ, Larson EB. Implementing the learning health system: from concept to action. *Ann Intern Med*. 2012;157(3):207-10.
7. Richesson R.L., Smerek M. Electronic Health Records-Based Phenotyping. In: Collaboratory NR, editor. *Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials*. Durham, NC: Duke Clinical Research Institute; 2014.
8. Richesson RL, Hammond WE, Nahm M, Wixted D, Simon GE, Robinson JG, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. *J Am Med Inform Assoc*. 2013;20(e2):e226-31.
9. Collaboratory. NHCSR. *Rethinking Clinical Trials. A Living Textbook of Pragmatic Clinical Trials*. Durham: Duke University; 2015.
10. Rusincovitch SA. Practical Development and Implementation of EHR Phenotypes. Presented on NIH Collaboratory Grand Rounds, November 15, 2013. 2013. Available from: <https://www.nihcollaboratory.org/Pages/Grand-Rounds-11-15-13.aspx>.
11. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc*. 2014;21(4):578-82.
12. Nichols GA, Desai J, Elston Lafata J, Lawrence JM, O'Connor PJ, Pathak RD, 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. *Preventing chronic disease*. 2012;9:E110.
13. Prieto-Centurion V, Rolle AJ, Au DH, Carson SS, Henderson AG, Lee TA, et al. Multicenter study comparing case definitions used to identify patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2014;190(9):989-95.
14. Curtis LH, Weiner MG, Boudreau DM, Cooper WO, Daniel GW, Nair VP, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 1:23-31.
15. Granger BB, Staton M, Peterson L, Rusincovitch SA. Prevalence and Access of Secondary Source Medication Data: Evaluation of the Southeastern Diabetes Initiative (SEDI). *AMIA Jt Summits Transl Sci Proc*. 2015;2015:66-70.
16. Gregg EW, Li Y, Wang J, Burrows NR, Ali MK, Rolka D, et al. Changes in diabetes-related complications in the United States, 1990-2010. *N Engl J Med*. 2014;370(16):1514-23.
17. Mo H, Thompson WK, Rasmussen LV, Pacheco JA, Jiang G, Kiefer R, et al. Desiderata for computable representations of electronic health records-driven phenotype algorithms. *J Am Med Inform Assoc*. 2015.
18. Sun J, McNaughton CD, Zhang P, Perer A, Gkoulalas-Divanis A, Denny JC, et al. Predicting changes in hypertension control using electronic health records from a chronic disease management program. *J Am Med Inform Assoc*. 2014;21(2):337-44.
19. ONC. Interoperability Standards Advisory (ISA): Office of the National Coordinator for Health IT; 2016 [updated January 19, 2016; cited 2016 April 4]. Available from: <https://www.healthit.gov/standards-advisory>.
20. Zozul MN, Hammond WE, Green BB, Kahn MG, Richesson RL, Rusincovitch SA, et al. Assessing Data Quality for Healthcare Systems Data Used in Clinical Research (Version 1.0). An NIH Health Care Systems Research Collaboratory Phenotypes, Data Standards, and Data Quality Core White Paper. 2013 [cited 2015 June 22]. Available from: <https://sites.duke.edu/rethinkingclinicaltrials/tag/data-quality/>.
21. Holve E, Kahn M, Nahm M, Ryan P, Weiskopf N. A comprehensive framework for data quality assessment in CER. *AMIA Jt Summits Transl Sci Proc*. 2013;2013:86-8.
22. Kahn MG, Brown JS, Chun AT, Davidson BN, Meeker D, Ryan PB, et al. Transparent Reporting of Data Quality in Distributed Data Networks. *eGEMs (Generating Evidence & Methods to improve patient outcomes)* 2015;3(1):Article 7.
23. CMS. Quality Measure Development and Management Overview: Centers for Medicare & Medicaid Services; 2016 [cited 2016 April 11]. Available from: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Quality-Measure-Development-Lifecycle-Overview.pdf>.
24. University. V. PheKB 2012 [cited 2013 May 24]. Available from: <http://www.phekb.org/>.
25. Xu J, Rasmussen LV, Shaw PL, Jiang G, Kiefer RC, Mo H, et al. Review and evaluation of electronic health records-driven phenotype algorithm authoring tools for clinical and translational research. *J Am Med Inform Assoc*. 2015.
26. PheMA. PheMA wiki: Phenotype Execution Modeling Architecture project Mayo Clinic; 2015 [cited 2015 September 28]. Available from: http://informatics.mayo.edu/phema/index.php/Main_Page.
27. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform*. 2015;216:574-8.
28. Mini-Sentinel. Distributed Database and Common Data Model: Mini-Sentinel Coordinating Center; 2015 [updated 1/9/2015; cited 2015 Aug 30]. Available from: http://mini-sentinel.org/data_activities/distributed_db_and_data/details.aspx?ID=105.



29. OHSDI. OMOP Common Data Model 2015 [cited 2015 August 30]. Available from: <http://www.ohdsi.org/data-standardization/the-common-data-model/>.
30. PCORnet. PCORnet Common Data Model (CDM). Why, What, and How? : PCORnet Coordinating Center; 2015 [cited 2015 Aug 30]. Available from: <http://www.pcornet.org/pcornet-common-data-model/>.
31. Jiang G, Solbrig HR, Kiefer R, Rasmussen LV, Mo H, Speltz P, et al. A Standards-based Semantic Metadata Repository to Support EHR-driven Phenotype Authoring and Execution. *Stud Health Technol Inform*. 2015;216:1098.
32. Richesson RL. An Informatics Framework for the Standardized Collection and Analysis of Medication Data in Networked Research: *Journal of Biomedical Informatics* - in press; 2014.
33. OHSDI. CALYPSO (Criteria Assessment Logic for Your Population Study in Observational data): *Observational Health Data Science and Informatics*; 2016 [cited 2016 April 11]. Available from: <http://www.ohdsi.org/analytic-tools/calypso-for-study-population-evaluation/>.
34. OHSDI. Criteria Assessment Logic for Your Population Study in Observational data (CALYPSO): *Observational Health Data Sciences and Informatics*; 2015 [updated August 8, 2015; cited 2015 September 28]. Available from: <http://www.ohdsi.org/analytic-tools/calypso-for-study-population-evaluation/>.
35. Wilcox A, Randhawa G, Embi P, Cao H, Kuperman G. Sustainability Considerations for Health Research and Analytic Data Infrastructures. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 2014;2(2):Article 8.
36. Bozeman B, Baoardman C. An Evidence-Based Assessment of Research Collaboration and Team Science: Patterns in Industry and University-Industry Partnerships. Paper commissioned for the National Research Council Study of the Science of Team Science. Workshop on Institutional and Organizational Supports for Team Science; October 24; Washington, D.C.: The National Academies of Sciences, Engineering, and Medicine; 2013.
37. Papadaki M, Hirsch G. Curing consortium fatigue. *Sci Transl Med*. 2013;5(200):200fs35.