

Research and Applications

GARDE: a standards-based clinical decision support platform for identifying population health management cohorts

Richard L. Bradshaw ^{1,2}, Kensaku Kawamoto ^{1,2,*}, Kimberly A. Kaphingst^{3,4}, Wendy K. Kohlmann^{2,3,5}, Rachel Hess^{2,5,6}, Michael C. Flynn^{2,6,7}, Claude J. Nanjo^{1,2}, Phillip B. Warner^{1,2}, Jianlin Shi¹, Keaton Morgan ^{1,2,8}, Kady Kimball³, Pallavi Ranade-Kharkar^{1,9}, Ophira Ginsburg¹⁰, Melody Goodman¹¹, Rachelle Chambers¹⁰, Devin Mann¹⁰, Scott P. Narus¹, Javier Gonzalez¹⁰, Shane Loomis^{10,12}, Priscilla Chan¹⁰, Rachel Monahan¹⁰, Emerson P. Borsato¹, David E. Shields^{1,2}, Douglas K. Martin^{1,2}, Cecilia M. Kessler^{2,4}, and Guilherme Del Fiol ^{1,*}

¹Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah, USA, ²University of Utah Health, Salt Lake City, Utah, USA, ³Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA, ⁴Department of Communication, University of Utah, Salt Lake City, Utah, USA, ⁵Departments of Population Health Sciences, University of Utah, Salt Lake City, Utah, USA, ⁶Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA, ⁷Community Physicians Group, University of Utah, Salt Lake City, Utah, USA, ⁸Department of Surgery, University of Utah, Salt Lake City, Utah, USA, ⁹Intermountain Healthcare, Salt Lake City, Utah, USA, ¹⁰New York University Langone Health, New York City, New York, USA, ¹¹School of Global and Public Health, New York University, New York City, New York, USA, and ¹²Epic Systems Corporation, Madison, Wisconsin, USA

Corresponding Author: Richard L. Bradshaw, MS, PhD, Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Suite 140, Salt Lake City, UT 84108-3514, USA; rick.bradshaw@utah.edu

*Co-senior author.

Received 10 December 2021; Revised 3 February 2022; Editorial Decision 10 February 2022; Accepted 18 February 2022

ABSTRACT

Population health management (PHM) is an important approach to promote wellness and deliver health care to targeted individuals who meet criteria for preventive measures or treatment. A critical component for any PHM program is a data analytics platform that can target those eligible individuals.

Objective: The aim of this study was to design and implement a scalable standards-based clinical decision support (CDS) approach to identify patient cohorts for PHM and maximize opportunities for multi-site dissemination.

Materials and Methods: An architecture was established to support bidirectional data exchanges between heterogeneous electronic health record (EHR) data sources, PHM systems, and CDS components. HL7 Fast Healthcare Interoperability Resources and CDS Hooks were used to facilitate interoperability and dissemination. The approach was validated by deploying the platform at multiple sites to identify patients who meet the criteria for genetic evaluation of familial cancer.

Results: The Genetic Cancer Risk Detector (GARDE) platform was created and is comprised of four components: (1) an open-source CDS Hooks server for computing patient eligibility for PHM cohorts, (2) an open-source Population Coordinator that processes GARDE requests and communicates results to a PHM system, (3) an EHR Patient Data Repository, and (4) EHR PHM Tools to manage patients and perform outreach functions. Site-specific

deployments were performed on onsite virtual machines and cloud-based Amazon Web Services.

Discussion: GARDE's component architecture establishes generalizable standards-based methods for computing PHM cohorts. Replicating deployments using one of the established deployment methods requires minimal local customization. Most of the deployment effort was related to obtaining site-specific information technology governance approvals.

Key words: clinical decision support system (D020000), population health management (D000076602), Health Level Seven (D057208), FHIR, CDS Hooks

INTRODUCTION

Population health management (PHM) is an important approach to promote wellness and deliver health care by targeting specific subgroups of individuals who meet criteria for certain preventive or treatment measures.¹ Several initiatives are promoting the adoption of PHM practices to improve health. For example, the Population Health Program Accreditation Standards, set by the US National Committee for Quality Assurance (NCQA), have been defined to evaluate healthcare organizations' programs against best practices for PHM.² A critical component for any PHM program is a data analytics platform with algorithms that can automatically screen and identify patient cohorts who are eligible to receive the target interventions.³

Given the complexity of algorithms needed to identify target patient cohorts,⁴ sharing of algorithms among organizations is highly desirable. While most market-leading EHR vendors provide tools for identifying PHM cohorts, in general, the logic is difficult to share with other organizations and the cohort selection logic has limitations. For example, in a related study, we augmented rule-based logic in Genetic Cancer Risk Detector (GARDE) with information extracted via a natural language processing service that was not otherwise available through the EHR to increase algorithm accuracy.⁵ In addition, EHR-based analytical capabilities are generally proprietary and noninteroperable, making them difficult to share between organizations that use different EHR platforms.⁶ Standards-based methods are needed to scale beyond single-vendor solutions and to reduce the implementation barriers related to disseminating patient cohort identification logic. However, the NCQA PHM program does not specify standards for interoperability with various patient data sources such as EHR systems, health information exchanges, and patient registries.

While there has been significant work on methods that support scaling and sharing interoperable *patient-level* clinical decision support (CDS) tools,^{7–14} very little has been published on interoperable population-level CDS methods to be applied across a patient population or healthcare system.¹⁵ Kukhareva et al described a system that computes national quality measures for a healthcare organization using standards-based open-source CDS methods, but the solution was designed to compute quality measures as opposed to identify target cohorts. Other relevant population-based methods have been investigated to support clinical trial cohort selection and recruitment,^{16–18} and each described toolsets for creating cohort queries. Interoperability was addressed by normalizing heterogeneous EHR data to a common model based on the Health Level Seven (HL7) version 3 Reference Information Model¹⁹ or OpenEHR archetypes.²⁰ However, HL7 has been reducing its support for version 3, favoring focus on newer approaches based on the Fast Healthcare Interoperability Resources (FHIR) standard, which is rapidly being adopted by most large vendor-based EHRs.²¹ In addition,

these previous approaches did not provide a standards-based mechanism for sharing algorithm logic.

To address interoperability and scalability challenges in support of PHM, we created GARDE, a novel open-source and EHR-agnostic population CDS approach leveraging emerging standards FHIR and CDS Hooks.²² While several studies have used CDS Hooks to implement patient-level CDS tools,¹⁴ to our knowledge, GARDE is the first attempt to use FHIR and CDS Hooks as a generalizable method for population-level CDS. We describe GARDE's (1) architecture; (2) deployment approaches; and (3) multi-site validation using GARDE to identify patients who met guideline-based criteria for genetic evaluation of familial cancer in a multi-site randomized controlled trial.^{23,24}

MATERIALS AND METHODS

Architecture

The design of GARDE needed to support (1) reading patient facts from EHR data sources to evaluate patients for cohort eligibility via CDS logic; (2) communicating patients who met eligibility criteria to the PHM system, and (3) performing these operations at multiple sites. From an architectural perspective, these requirements implied GARDE needed to support bidirectional data exchanges between heterogeneous EHR data sources, PHM systems, and CDS components. The emerging CDS interoperability standards of FHIR and CDS Hooks were selected to facilitate dissemination.

Interoperability standards

HL7 FHIR was selected for data representation and exchange using FHIR application programming interfaces (APIs) and *Resources*. FHIR *Resources* are entities (data models) for exchanging healthcare data, with *Resources* such as *Patient*, *Practitioner*, *Condition*, *Observation*, and *Family Member History*.²⁵ *Resources* are comprised of structured data elements with bindings to standard terminologies for coded data. FHIR APIs are based on RESTful Web Services²⁶ used for requesting/sending FHIR *Resources* between systems.

HL7 CDS Hooks allows client applications to submit patients for evaluation. The CDS Hooks specification defines a set of "hooks" that are triggered by specific user events in the EHR, such as opening a patient's chart (*patient-view hook*) or selecting a medication for ordering (*order-select hook*). When a *hook* is triggered, a CDS Hooks request is sent to designated CDS services. Requests may include patient data in FHIR format according to data requirements specified in the CDS service's *prefetch* signature. In addition, the CDS request may provide an access token for the CDS service to access the EHR's FHIR server on-demand. CDS services then use CDS logic to evaluate the patient and make a conclusion (eg, assessment, suggestion) that is sent back to the EHR in the form of CDS Hooks response "cards."

For example, while a clinician is prescribing a medication, the EHR's *order-select* hook could trigger a CDS service that identifies and offers less expensive medications during the ordering process. As this example illustrates, however, CDS Hooks requests are typically triggered by user interactions with the EHR in the context of individual patients. We investigated an approach that uses CDS Hooks to process populations of patients asynchronously outside the context of EHR use. To support population-level processing, we added the equivalent of a time-triggered hook extension to the CDS Hooks specification. Unlike hooks that are triggered upon certain user events within an EHR, the time-triggered event invokes asynchronous population-based evaluations at predetermined times.

The CDS Hooks specification does not mandate a standards-based rules/logic formalism.^{27,28} Instead, CDS Hooks achieves interoperability by standardizing how it interfaces with EHRs (ie, how it is triggered, how it obtains input data, and how it returns CDS responses). The only requirement for CDS rules is to consume data based on FHIR. While this architecture supports the use of standards-based logic formats such as the HL7 Clinical Quality Language,²⁸ any knowledge representation approach can be used in the proposed solution.

Building blocks

Architects, informaticists, and engineers identified architectural building blocks for GARDE: (1) a patient data warehouse to select relevant target populations through queries; (2) a standards-based CDS server to evaluate patient cohort eligibility; (3) EHR/vendor-provided data services to communicate identified cohorts to the EHR's PHM system; (4) EHR/vendor-provided PHM tools to manage identified cohorts (review patients, conduct outreach, etc.); (5) a component to extract, transform, and load (ETL) data from multiple data sources and sites to normalize and prepare incoming data for analysis; and (6) a component to manage service choreography. The service choreography component would direct population extractions and CDS results back to the PHM system using ETL tools and interoperability standards to mediate component communication heterogeneity.

Deployment strategies

To support wide adoption, GARDE needed to be easily deployed at different healthcare organizations with various network topologies, server configurations, security policies, EHRs, and information technology (IT) expertise. Deployment options needed to be flexible and configurable. The "highly scalable" requirement also implied GARDE needed to perform adequately moving, processing, and computing CDS logic for enterprise-scale populations. Deployment strategies were designed and implemented based on a case study and are described in the results.

Multi-institutional validation

GARDE was deployed in three large healthcare organizations—University of Utah Health (UHealth), New York University Langone Health (NYU), and Intermountain Healthcare—with two different EHRs (Epic® and Cerner®) and different IT infrastructures/policies. Two of the deployments, UHealth and NYU (Epic® sites), were in support of the Broadening the Reach, Impact, and Delivery of Genetic Evaluation (BRIDGE) trial, a multi-site randomized controlled trial aimed at comparing two genetic services delivery models for patients meeting criteria for genetic evaluation for hereditary cancers.²² The third deployment was at Intermountain Healthcare, a re-

search partner who collaborated with testing GARDE interoperability with Cerner®. For all three deployments, GARDE used the same logic to identify patients who meet criteria set by National Comprehensive Cancer Network (NCCN) guidelines for genetic testing for hereditary cancers using patients' family health history in the EHR. Details about the CDS logic and integration with clinical workflow are described elsewhere.²³

Site-specific deployment requirements

UHealth

UHealth served as the home institution for the development and initial piloting of GARDE. For healthcare software deployment and access to data sources, UHealth uses virtual machines (VMs) inside the organization's network with direct access to the institution's enterprise data warehouse (EDW) EHR schemas, an additional EDW schema to store application data, and direct access to Epic data interfaces/services for loading platform output into Epic's PHM system.

NYU

NYU required GARDE to be deployed on cloud-based Amazon Web Services (AWS) to avoid having external software inside their healthcare network with direct access to patient data sources.

Intermountain

Intermountain had a hybrid set of requirements between those at UHealth and NYU. External software could be installed on Intermountain VMs but could not directly access Intermountain's patient data sources.

In addition to differences in IT requirements and environments, each site used different terminology codes to represent family health history, even when using the same EHR. Therefore, mappings between local data models and local codes to FHIR and standard terminologies were needed at all three sites.

RESULTS

Architecture

Overall, the architecture contains four components: *OpenCDS*, *Population Coordinator*, *EHR Patient Data Repository*, and *EHR PHM Tools* (Figure 1). *OpenCDS* is an open-source CDS Hooks-compliant server that computes patient eligibility for PHM cohorts. *Population Coordinator* is the application endpoint and service choreographer that receives platform requests, processes population data (transformations to/from FHIR), performs population-based CDS communications, and sends results to the PHM system. *EHR Patient Data Repository* is the source for patient data used by the CDS logic. *EHR PHM Tools* include a registry where patients who met PHM criteria are tracked and a dashboard clinical staff use to navigate the registry, review individual patient data, and perform patient outreach functions.

Population Coordinator choreography

The *Population Coordinator* may be invoked as often as required by the PHM application. For the current installations, population evaluation invocations are triggered by scheduled jobs nightly and weekly. The CDS Hooks *hook* in this case is a scheduled event that invokes a population-based evaluation rather than an individual patient evaluation. A method was added to the open-source OpenCDS CDS Hooks services to handle population evaluations.

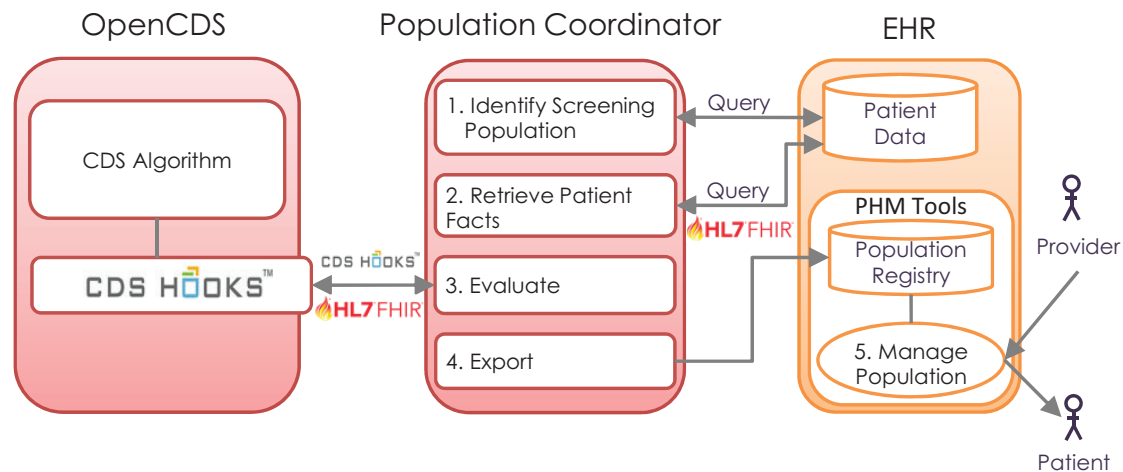


Figure 1. GARDE's component architecture diagram. GARDE, Genetic Cancer Risk Detector.

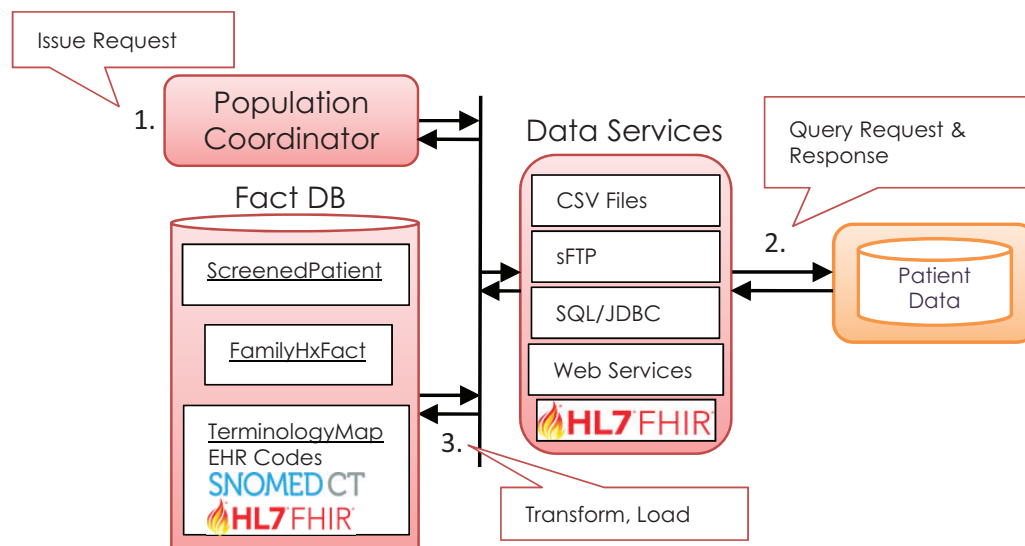


Figure 2. Population Coordinator's workflow to identify, extract, transform, and load patient data from a patient database into the FactDB. This is a detailed view of the process required to perform steps 1 and 2 from Figure 1.

The following sections describe GARDE's service choreographies following the numeric steps in Figure 1.

Step 1—identify screening population

The goal of step 1 is to identify the screening population, or the population GARDE will evaluate. While the screening population could be “every patient in the database,” step 1 is an opportunity to preprocess and filter out unnecessary data using an organization-specific database query/logic before the more complex CDS operations are computed. Figure 2 is a detailed diagram of the process.

The *Population Coordinator* executes an ETL pattern to identify and retrieve the screening population. In Figure 2 process 1, an *Extract Population* request is issued to the *Population Coordinator* with a parameter specifying which screening population to extract. In process 2, the *Population Coordinator* retrieves the requested population using a data service designated by configuration. In process 3, the retrieved population is transformed and loaded into the *FactDB* in preparation for the processes to follow.

The *Data Service* layer in Figure 2 is responsible for handling inbound and outbound data requests from multiple common data source technologies. The *FactDB* is the central data store that serves multiple purposes: (1) provides a persistence mechanism for GARDE for tracking and managing patient cohorts, patient facts, and data provenance, (2) supports interoperability by using FHIR data elements and terminology, and (3) serves as a staging area for intermediate data to improve performance.²⁹

Step 2—retrieve patient facts

The goal of step 2 is to collect and normalize patient facts required by the CDS algorithm. Patient facts are retrieved following the ETL pattern described in step 1 with the following variations:

- Process 1—the issued request specifies which patient facts to retrieve (versus specifying which screening population);
- Process 2—fact-specific data services and queries are selected to retrieve fact data; and

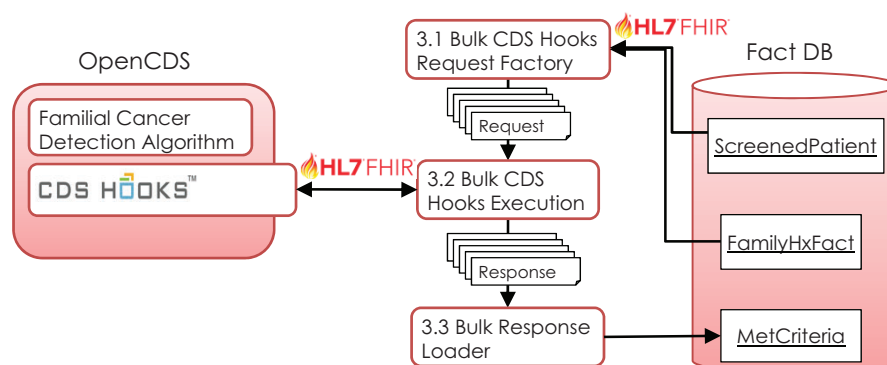


Figure 3. Population Coordinator bulk CDS evaluation process; a detailed view of [Figure 1](#) step 3. CDS: clinical decision support.

- Process 3—data transforms to FHIR (versions STU3 or R4) are based on the facts requested, and the *FactDB* data model used to store data is based on the FHIR type of the requested facts.

Other than these variations, the fact retrieval process follows the same pattern as the screening population retrieval process.

Step 3—evaluate

After all patient facts are retrieved and persisted, the *Population Coordinator* prepares and executes CDS evaluations for the designated population ([Figure 3](#)).

A *Population Coordinator* request to evaluate a population includes variables that designate the population and CDS evaluation module. The request then invokes operations 3.1 through 3.3 from [Figure 3](#). Operation 3.1 builds all CDS Hooks requests. CDS module-specific patient facts are extracted from the *FactDB* in bulk and serialized into CDS Hooks requests with FHIR *Resources* using a *Bulk CDS Hooks Request Factory*. Operation 3.2 can send CDS requests in rapid succession patient-by-patient (following the CDS Hooks specification) or using a bulk operation that we added to the CDS Hooks server as an extension of the base specification. Upon CDS evaluation completion, a bulk CDS Hooks response is returned to the *Population Coordinator* either patient-by-patient or in bulk based on how it was called, and in operation 3.3, responses are interpreted, transformed into *FactDB* artifacts, and loaded into the *FactDB* where they are accessible for downstream processes.

The bulk methods described for operation 3.2 were not part of the CDS Hooks specification; they were added based on prior work to improve performance.²⁹ A new *hook* was added as well, *scheduled-population-evaluation*, to support population CDS workflows. An informal study was performed to estimate the performance impact ([Table 1](#)).

Step 4—Export

To populate the PHM system with CDS results, results are communicated using any of the supported data exchange mechanisms. Communications are outgoing messages facilitated by the export process following the same *Population Coordinator* patterns ([Figure 4](#)).

In [Figure 4](#) process 1, a request to export a specific population is issued to the *Population Coordinator*. Process 2 interprets the request and uses the designated *Data Service* to export *MetCriteria* facts from GARDE. Process 3 transforms *MetCriteria* into the data structure required by the PHM registry loader and invokes the import service.

Table 1. Performance metrics comparing patient by patient request processing versus processing populations using bulk methods

Step	Patient by patient (pats/s)	In bulk (pats/s)
3.1 Compose CDS requests	0.75	832
3.2 CDS request execution	61	2566
3.3 CDS response loading	(not available)	1102

CDS: clinical decision support.

Step 5—manage population

At UHealth and NYU, Epic's Healthy Planet feature set was used for PHM ([Figure 5](#)). In [Figure 5](#), cohort patients are displayed in the top main panel. Since lists are often very large, *Filters* are provided to find specific subpopulations. The bottom half of the report is dedicated to the selected patient's clinical summary configured specifically for the given population. In this case, this clinical summary provides information on family history, genetic counseling encounters, and outreach status.

Outreach functions are accessed by selecting a patient or group of patients from the list, selecting an outreach message from a list of message templates, and then clicking the send option. Personalized messages are sent to each outreach recipient through the patient portal.

Deployment strategies

Deployment strategies resulted from the participant's site requirements. The GARDE components that needed to be deployed were the *Population Coordinator*, *FactDB*, and *OpenCDS*. Three deployment hosting strategies are supported:

1. On premises deployment — GARDE components are installed on the implementing site's local servers, typically VMs.
2. Cloud deployment — GARDE components are installed on a cloud-based solution. Current cloud solutions include AWS, Utah's Center for High Performance Computing services, and Azure in 2022.
3. Hybrid deployment — some GARDE components are deployed on premises VMs and others are deployed on a cloud-based solution.

After determining deployment strategy, strategies to import patient facts and export results to the PHM system were determined based on *Population Coordinator Data Service* capabilities. Two patient fact import strategies were implemented: (1) via database access and (2) via secure structured text file sharing. Two population

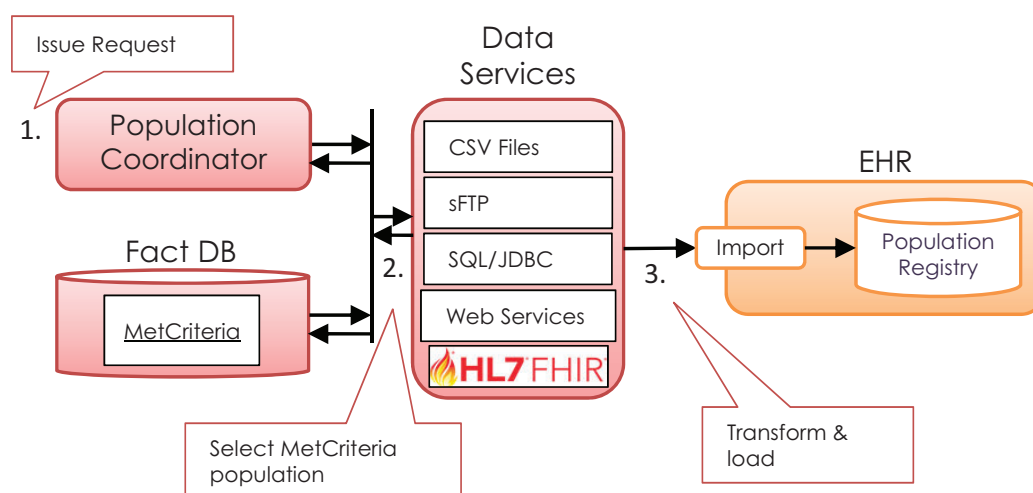


Figure 4. The Population Coordinator's population export process writing to the EHR's import services. HER: electronic health record.

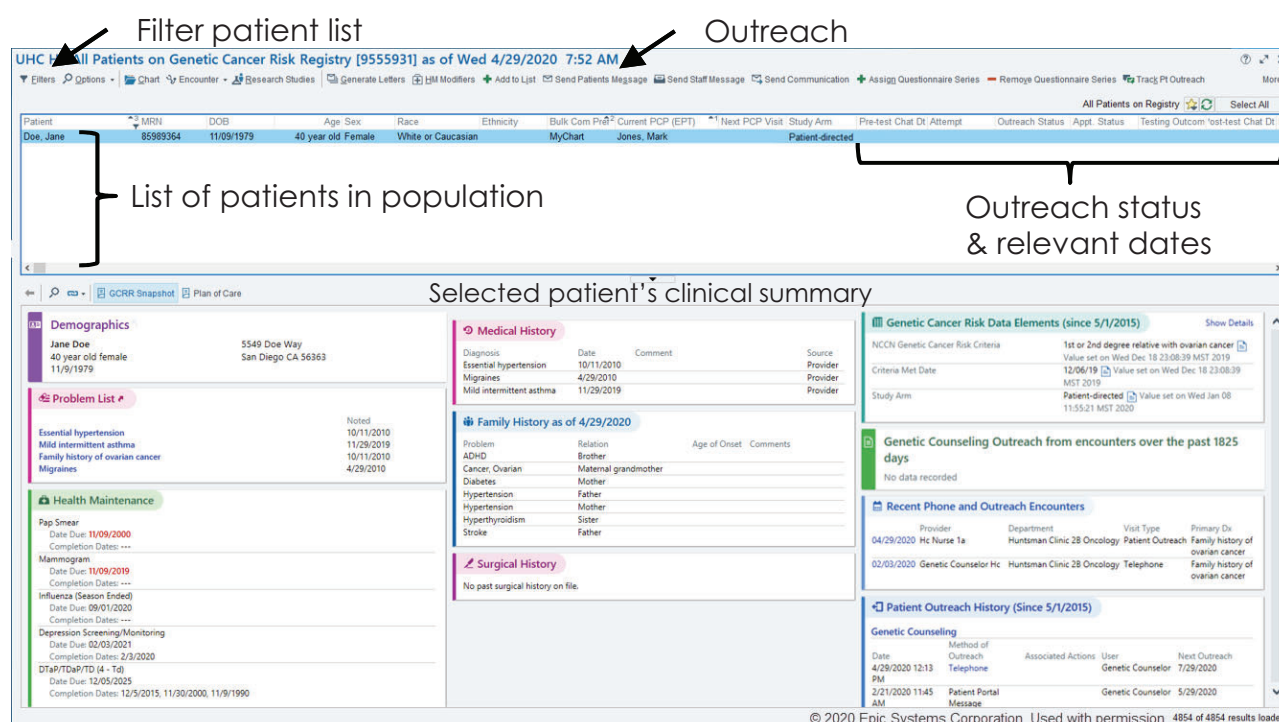


Figure 5. Population health management application based on Epic's Healthy Planet.

export-to-PHM system strategies were implemented: (1) via secure structured text file sharing and (2) via EHR Web services APIs.

Terminology mappings are required for sites that do not support terminologies used by GARDE's CDS logic. Mappings between the implementing site's codes and standard terminologies are conducted by data analysts for each deployment site using tools provided by our team. The completed mappings are loaded into the *FactDB* terminology tables where they are used for ETL processes.

Multi-site validation

GARDE was successfully deployed at each of the validation sites. Descriptions of each deployment are as follows.

UHealth

GARDE was deployed on site by the Utah project team. Patient data are accessed directly from the EDW and processed on a local CDS Hooks server, and results are written directly to the EHR's PHM system. GARDE identified 135 817 screening patients; 5155 (3.8%) met NCCN criteria for genetic testing.

NYU

GARDE was deployed on an NYU-managed AWS cloud environment. Utah engineers deployed all components on the designated cloud. Patient data are exchanged via de-identified structured text files prepared and saved to an AWS file share by NYU staff. The saved files trigger the cloud-based platform that responds by output-

ting structured text files back to the file share. NYU processes detect and send the output files to an interface engine that writes results to their PHM system. At NYU, GARDE identified 208 071 screening patients; 14 200 (6.8%) met NCCN criteria for genetic testing.

Intermountain

GARDE's CDS logic was deployed locally with direct patient database access to assess interoperability with the Cerner EHR. The deployed logic identified ~741 000 screening patients, ~21 500 (2.9%) of whom met NCCN criteria for genetic testing. Intermountain is currently investigating PHM solutions and plans to deploy the full GARDE platform locally.

A randomized controlled trial is underway with 2780 patients at UHealth and NYU to compare genetic counseling outreach with an automated conversational agent (chatbot) that provides patients with familial cancer education and offers genetic testing to patients who met NCCN criteria for genetic testing to standard of care (1U01CA232826—Kaphingst and Sigireddi MPis).

DISCUSSION

Architecture

GARDE establishes a generalizable EHR integration architecture that facilitates incorporating cutting edge analytics to identify cohorts for PHM using existing HL7 standards, that is, CDS Hooks and FHIR. CDS Hooks provides a standard mechanism to specify data requirements for a specific CDS service, a standard mechanism to exchange the required data, and a standard mechanism for applications to communicate with the CDS service. In the case of GARDE, the resulting CDS Hooks output is a card with a *CommunicationRequest* (FHIR) that communicates the outcome, for example, that the patient meets NCCN criteria for genetic testing. Using this strategy, GARDE algorithms are EHR agnostic and are not bound to the capabilities of a specific EHR. Algorithms are accessed using an interoperable approach, maximizing the sharing of GARDE logic and population analysis capabilities across EHR products and healthcare organizations.

The current version of GARDE focuses on cohort identification utilizing the native EHR's PHM tool for functions such as outreach tracking, bulk orders, bulk messaging, and documentation. This decision was made as many fully featured EHRs such as Epic already have a robust PHM tool that is tightly integrated with other elements of the EHR and is incorporated into existing user workflows. In other words, replacing these aspects of the EHR's PHM tool would not have added functionality and may have actually decreased usability. At the same time, there may be EHR platforms—such as some EHRs that may be used in low-resourced healthcare settings—that do not have a robust PHM tool natively available. Thus, a valuable potential future direction of this work will be to develop an open-source, standards-based PHM tool that provides these capabilities for healthcare settings that do not currently have access to such functionality.

The central component of GARDE, the *Population Coordinator*, is the component that facilitates performant, interoperable, population-based communications between heterogeneous EHR data sources, OpenCDS, and PHM tools. CDS Hooks and FHIR standards provide the architectural blueprints for interoperability, in which CDS rules and services in OpenCDS can be integrated with GARDE through CDS Hooks requiring minimal local customization.

Population Coordinator Data Services were designed specifically to create semantically consistent FHIR using terminology services and ETL strategies to resolve EHR data inconsistencies. Consequently, GARDE successfully maintains internal interoperability between components and consistency across installations.

The *Data Services* were performant. Bulk processes, versus patient-by-patient processes, improved performance by two orders of magnitude following similar practices described elsewhere.²⁹ Additional performance improvements are possible by turning off data provenance persistence. This would be an attractive option when GARDE is used as a service-level component where other components manage data provenance and analytics.

A potential future direction will be to use FHIR Bulk Data Access³⁰ for extracting relevant FHIR in an EHR-independent and scalable manner. The extracted data then could be used by GARDE for data analysis and cohort identification. Currently, the FHIR Bulk Data Access standard has some limitations, such as not yet being available for production use in many EHR platforms and having limited capabilities for restricting the scope of the extracted resources (eg, an inability to limit the extraction to specified lab results as opposed to all lab results). Nevertheless, with increasing capabilities being specified and emerging EHR vendor support underway,³¹ we plan to explore supporting FHIR Bulk Data Access in an upcoming release.

Deployment strategies

The offered deployment strategies support different requirements from the three case study deployments. The resulting deployment options are versatile, permitting deployment in personal computers, enterprise VMs, or auto-scaling cloud environments. The component-based architecture and flexible configuration options also allow deploying individual components in any combination of these environments. This modular approach has several advantages, including flexibility to support various IT architectures and data privacy policies, as well as the ability to isolate and debug individual components.

Most of the effort and time required for each deployment was related to obtaining site-specific IT governance approvals. Once governance approvals had been obtained, system integration tasks, such as mapping terminology codes between local codes and standard codes, were required to achieve interoperability and deployment.

Multi-site validation

The multi-site validation study provided the requirements for GARDE development and implementation. GARDE met the validation study's requirements by identifying over 42 000 patients from the three sites. Over 19 000 of the identified patients were placed in operational PHM registries that genetic counseling staff are actively using to manage patients at UHealth and NYU. Algorithm results and underlying population disparities between the reported sites are being evaluated and reported through the BRIDGE trial.²⁴

Limitations

GARDE has limitations. First, GARDE was designed to support multiple PHM use cases, but the familial cancer risk case study is the only operational implementation supported to date. We are currently in the planning stages of adding more cancer-related use cases to enhance GARDE's generalizability and scalability. Second, GARDE is highly configurable and requires in-depth knowledge of the system components to deploy and run. Deploying GARDE with-

out assistance is possible but challenging and we are currently working on simplifying the deployment process. To achieve true plug-and-play interoperability, EHR vendors will need to adopt and enforce data standards more rigorously down to the level of individual data elements, potentially following tighter specifications such as in FHIR profiles. Last, this paper only covers technical approaches for platform interoperability and deployment. GARDE implementation at the study sites involved a sociotechnical process that is typical of complex health IT implementations, including steps such as clinical prioritization, IT governance approvals, stakeholder education, and user training.

CONCLUSION

The architecture and deployment strategies for GARDE, a novel open-source and system-agnostic population CDS platform based on FHIR and CDS Hooks, have been described. Multi-site validation demonstrated GARDE's standards-based design met scalability and interoperability requirements for three institutions identifying over 42 000 patients who met criteria for genetic testing of familial cancers based on their family health history in the EHR, 19 000 of whom have been placed into operational PHM registries used daily by genetic counselors and providers for patient outreach and management. Future work includes adding support for the FHIR Bulk Data Access specification, expanding GARDE to support other use cases, simplifying the deployment process, and deployment at other institutions. Access to the open-source software is available from <https://bitbucket.org/RickSlc/garde-app>. The OpenCDS source is available for free by registering for an account at <https://www.opencds.org/pages/join>.

FUNDING

This work was supported by Grant No. U24CA204800 from the Informatics Technology for Cancer Research program of the US National Cancer Institute and by the National Cancer Institute of the National Institutes of Health, USA, under award number U01CA232826.

AUTHOR CONTRIBUTIONS

Each author made substantial contributions to (1) the conception or design of the work (all authors); (2) the acquisition, analysis, or interpretation of data (RLB, KenK, KimK, WKK, RH, MCF, CJN, PBW, JS, KM, KadyNK, PR-K, OG, MG, RC, DM, SPN, JG, SL, PC, RM, and GDF); (3) the creation of new software used in the work (RLB, KenK, WKK, CJN, PBW, JS, KM, PR-K, SPN, JG, SL, EPB, DES, DKM, and GDF); or (4) the drafting or substantial revision of the manuscript (RLB, KenK, KimK, WKK, RH, CJN, JS, KM, PC, EPB, and GDF). All authors also approved the manuscript for submission and agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work are appropriately investigated, resolved, and documented in the literature.

CONFLICT OF INTEREST

KenK reports honoraria, consulting, sponsored research, licensing, or co-development in the past year with Hitachi, Pfizer, RTI International, NORC at the University of Chicago, the University of California at San Francisco, Indiana University, MD Aware, the Korean Society of Medical Informatics, and the U.S. Office of the National Coordinator for Health IT (via Security Risk Solutions) in the area of health IT. KenK was also an unpaid board member of the nonprofit Health Level Seven International health IT standard development organization, he is an unpaid member of the U.S. Health Information

Technology Advisory Committee, and he has helped develop a number of health IT tools, which have been or may be commercialized to enable wider impact. None of these relationships have direct relevance to the manuscript but are reported in the interest of full disclosure.

DATA AVAILABILITY STATEMENT

Access to the open-source software is available from <https://bitbucket.org/RickSlc/garde-app>. Access to OpenCDS open-source software is available for free by registering for an account at <https://www.opencds.org/pages/join>.

REFERENCES

- Horn DM, Haas JS. Covid-19 and the mandate to redefine preventive care. *N Engl J Med*. 2020; 383 (16): 1505–7.
- HEDIS and Performance Measurement ncqa.org: National Committee for Quality Assurance; Retrieved December 3, 2021, from <https://www.ncqa.org/hedis/>.
- ncqa.org. NCQA Release New Standards Category—Population Health Management: NCQA.org; Retrieved December 3, 2021 from <https://www.ncqa.org/news/ncqa-release-new-standards-category-population-health-management/>.
- Shivade C, Raghavan P, Fosler-Lussier E, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. *J Am Med Inform Assoc* 2014; 21 (2): 221–30.
- Mowery DL, Kawamoto K, Bradshaw R, et al., ed. *Determining Onset for Familial Breast and Colorectal Cancer from Family History Comments in the Electronic Health Record*. San Francisco, CA: AMIA Informatics Summit; 2019.
- Sittig DF, Wright A, Osheroff JA, et al. Grand challenges in clinical decision support. *J Biomed Inform* 2008; 41 (2): 387–92.
- Abedin Z, Hoerner R, Habboushe J, et al. Implementation of a Fast Healthcare Interoperability Resources-based clinical decision support tool for calculating CHA2DS2-VASc scores. *Circ Cardiovasc Qual Outcomes* 2020; 13 (2): e006286.
- Cho I, Kim J, Kim JH, Kim HY, Kim Y. Design and implementation of a standards-based interoperable clinical decision support architecture in the context of the Korean EHR. *Int J Med Inform* 2010; 79 (9): 611–22.
- Curran RL, Kukhareva PV, Taft T, et al. Integrated displays to improve chronic disease management in ambulatory care: a SMART on FHIR application informed by mixed-methods user testing. *J Am Med Inform Assoc* 2020; 27 (8): 1225–34.
- Dolin RH, Boxwala A, Shalaby J. A pharmacogenomics clinical decision support service based on FHIR and CDS Hooks. *Methods Inf Med* 2018; 57 (S 02): e115–e23.
- Goldberg HS, Paterno MD, Rocha BH, et al. A highly scalable, interoperable clinical decision support service. *J Am Med Inform Assoc* 2014; 21 (e1): e55–62–e62.
- Orenstein EW, Yun K, Warden C, et al. Development and dissemination of clinical decision support across institutions: standardization and sharing of refugee health screening modules. *J Am Med Inform Assoc* 2019; 26 (12): 1515–24.
- Zhang M, Velasco FT, Musser RC, Kawamoto K. Enabling cross-platform clinical decision support through Web-based decision support in commercial electronic health record systems: proposal and evaluation of initial prototype implementations. *AMIA Annu Symp Proc* 2013; 2013: 1558–67.
- Taber P, Radloff C, Del Fiore G, Staes C, Kawamoto K. New standards for clinical decision support: a survey of the state of implementation. *Yearb Med Inform* 2021; 30 (1): 159–71.
- Kukhareva PV, Kawamoto K, Shields DE, et al. Clinical Decision Support-based Quality Measurement (CDS-QM) framework: prototype implementation, evaluation, and future directions. *AMIA Annu Symp Proc* 2014; 2014: 825–34.
- Bucur A, van Leeuwen J, Chen NZ, et al. Cohort selection and management application leveraging standards-based semantic interoperability and a groovy DSL. *AMIA Jt Summits Transl Sci Proc* 2016; 2016: 25–32. [TQ3]

17. Fernandez-Breis JT, Maldonado JA, Marcos M, *et al.* Leveraging electronic healthcare record standards and semantic web technologies for the identification of patient cohorts. *J Am Med Inform Assoc* 2013; 20 (e2): e288-96-e296.
18. Marcos M, Maldonado JA, Martinez-Salvador B, Bosca D, Robles M. Interoperability of clinical decision-support systems and electronic health records using archetypes: a case study in clinical trial eligibility. *J Biomed Inform* 2013; 46 (4): 676-89.
19. HL7. HL7 Version 3: Reference Information Model (RIM). HL7 International; HL7; 2013.
20. openEHR opeEHR.org. Retrieved December 3, 2021 from <https://www.openehr.org>.
21. HL7. HL7 FHIR Release 4. hl7.org: HL7; 2018.
22. HL7 CDS Hooks 1.0: HL7 & Boston Children's Hospital; Retrieved December 3, 2021, from <https://cds-hooks.hl7.org/1.0/>.
23. Del Fiol G, Kohlmann W, Bradshaw RL, *et al.* Standards-based clinical decision support platform to manage patients who meet guideline-based criteria for genetic evaluation of familial cancer. *JCO Clin Cancer Inform* 2020; 4: 1-9.
24. Kaphingst KA, Kohlmann W, Chambers RL, *et al.*; BRIDGE research team. Comparing models of delivery for cancer genetics services among patients receiving primary care who meet criteria for genetic evaluation in two healthcare systems: BRIDGE randomized controlled trial. *BMC Health Serv Res* 2021; 21 (1): 542.
25. HL7. HL7 FHIR Resource List. hl7.org: HL7; 2021.
26. HL7. HL7 FHIR Services. hl7.org: HL7; 2021.
27. Spineth M, Rappelsberger A, Adlassnig KP. Implementing CDS Hooks communication in an Arden-Syntax-based clinical decision support platform. *Stud Health Technol Inform* 2018; 255: 165-9.
28. HL7. Clinical Quality Language (CQL). hl7.org: HL7; 2021.
29. Tippetts TJ, Warner PB, Kukhareva PV, Shields DE, Staes CJ, Kawamoto K. Challenges and solutions in optimizing execution performance of a Clinical Decision Support-Based Quality Measurement (CDS-QM) framework. *AMIA Annu Symp Proc* 2015; 2015: 1194-203.
30. HL7. FHIR Bulk Data Access (Flat FHIR). hl7.org: HL7; 2021.
31. Jones J, Gottlieb D, Mandel JC, *et al.* A landscape survey of planned SMART/HL7 bulk FHIR data access API implementations and tools. *J Am Med Inform Assoc* 2021; 28 (6): 1284-7.