Developing and implementing a heart failure data mart for research and quality improvement

Erin Abu-Rish Blakeney, Seth Wolpin, Danielle C. Lavallee, Todd Dardas, Richard Cheng & Brenda Zierler

Published online: 19 Apr 2018.

Submit your article to this journal ☇

View related articles ☇

View Crossmark data ☇

Taylor & Francis
Taylor & Francis Group

Check for updates

Developing and implementing a heart failure data mart for research and quality improvement

Erin Abu-Rish Blakeney[a], Seth Wolpin[a], Danielle C. Lavallee[b], Todd Dardas[c], Richard Cheng[c], and Brenda Zierler[a]

[a]Department of Biobehavioral Nursing and Health Informatics, University of Washington School of Nursing; [b]Department of Surgery, University of Washington School of Medicine; [c]Department of Medicine, University of Washington School of Medicine

ABSTRACT

Objective: The purpose of this project was to build and formatively evaluate a near-real time heart failure (HF) data mart. Heart Failure (HF) is a leading cause of hospital readmissions. Increased efforts to use data meaningfully may enable healthcare organizations to better evaluate effectiveness of care pathways and quality improvements, and to prospectively identify risk among HF patients. *Methods and procedures*: We followed a modified version of the Systems Development Life Cycle: 1) Conceptualization, 2) Requirements Analysis, 3) Iterative Development, and 4) Application Release. This foundational work reflects the first of a two-phase project. Phase two (in process) involves the implementation and evaluation of predictive analytics for clinical decision support. *Results*: We engaged stakeholders to build working definitions and established automated processes for creating an HF data mart containing actionable information for diverse audiences. As of December 2017, the data mart contains information from over 175,000 distinct patients and >100 variables from each of their nearly 300,000 visits. *Conclusion*: The HF data mart will be used to enhance care, assist in clinical decision-making, and improve overall quality of care. This model holds the potential to be scaled and generalized beyond the initial focus and setting.

Introduction

The pursuit of the "Triple Aim" of improving the patient care experience, improving the health of populations, and reducing the provision of costly, unnecessary, and inefficient healthcare requires a well-orchestrated system-level effort to transform healthcare (1). One strategy towards supporting this transformation is to leverage electronic health records (EHR) management (2). EHRs are a rich data source yet the design and functionality is often limited to support clinical and administrative operations thus they are relatively nascent in complex clinical decision support abilities. For example, embedded individualized risk scores are still a rarity in EHRs, as are population health management tools (2). A recent review by Terry and colleagues (3) noted the underdeveloped nature of the EHR citing the inability to handle complex reporting necessary to support clinical practice, population health or practice-based research. As Terry et al. (3) stated, "The full power of an EMR cannot be realized if this technology is designed and used simply as an electronic version of a paper record." (p. 705)

Increasingly, healthcare organizations are working towards creating data warehouses (DWs) that house data pulled from different clinical and administrative systems. In this manner, data analysis and reporting are more easily supported without affecting the transactional EHR system (4).

CONTACT Erin Abu-Rish Blakeney ✉ erin2@uw.edu 🖃 Department of Biobehavioral Nursing and Health Informatics, University of Washington School of Nursing, 1959 NE Pacific Street Box 357266 Seattle, WA 98195, USA.

However, even with DWs in place, service lines and quality improvement (QI) programs may be at a loss with how to consolidate and transform data into actionable information without intentional, multidisciplinary teams to guide the process.

One area where DWs offer an opportunity to support patient and population health is in Heart Failure (HF). HF is a significant public health problem, with an expected 46% increase in prevalence from 2012 to 2030. HF is anticipated to affect more than 8 million people in the United States (U.S.) by 2030 (5). There were approximately 1 million discharges for first-listed HF in 2010, a number expected to grow with the aging population (6). As the most frequent cause of readmissions to hospitals among all diagnoses for those age 65 and older, HF comprises 61% of readmissions within the first 15 days post-discharge among all hospital readmissions (6). The greatest risk for readmission is immediately following discharge and just before death (6).

HF patients experience dyspnea, fatigue, edema, cardiac arrhythmias, sleep disturbance, depression, anxiety, organ failure, and social isolation (because symptoms such as dyspnea make patients less likely to participate in social activities) (7). Clinical management of HF is complex and challenges providers because patient's health status is constantly changing and care is highly interdependent within and between interprofessional (IP) care teams (8). Like with most chronic illnesses, HF care is often fragmented which impacts care quality, safety, and other outcomes (9-14). Studies show that poor communication among IP health care teams is a major contributor to medical errors, which is now the 3rd leading cause of death in the United States (15-17).

"Real-time access to knowledge" in HF holds the potential to improve outcomes by allowing clinicians to tailor therapy and resources to an individualized level of risk. This strategy requires a strong analytics platform integrated with EHR data. Such a platform would support predictive analytics to proactively identify whether a patient has HF, even if not identified as a primary diagnosis, so that appropriate pathways can be implemented. From a quality perspective, both process and outcome measures are necessary to assess the impact of QI initiatives. Lastly, from a patient perspective, education about risk classification might help patients to be more proactive and engaged in managing their condition (18).

Objective

The objective of this project was to build and formatively evaluate a near-real time heart failure (HF) data mart. The purpose of this paper is to describe methods employed for developing this focused HF data mart so that other organizations can learn from our processes and findings. Data marts are similar to DWs but are more often organized to meet specific decision support needs such as a focus on specific clinical conditions (19,20). Data marts often pull from larger DWs and, as with the one described in this paper, can be characterized as a patient registry. Gliklich et al. (21) define a patient registry as "an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves a predetermined scientific, clinical, or policy purpose(s)" (p. 1).

Methods and procedures

Conceptual framework: Software Development Lifecycle

To provide a framework for our HF data mart project we utilized a modified version of the Software Development Lifecycle (SDLC) (Figure 1).

This framework provides structure to software development processes, ensuring that the software meets defined requirements and is well planned in the early stages of development; thereby, avoiding costly-reworking of code in later stages (22). Royce's foundational paper (23) about SDLCs has been used as an example of the 'waterfall method' which can be thought of as 'top down' programming. However, as Larman and Basili (24) point out in their review article, Royce actually intended for his
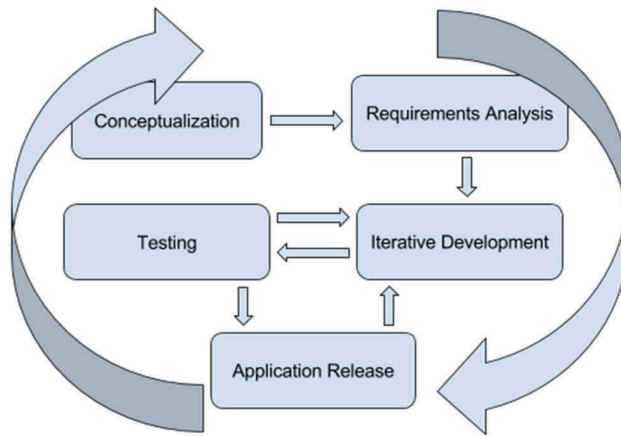
Figure 1. Modified software development lifecycle.

waterfall approach to include a second pass, allowing for iterative development. Recognizing that additional specifications and needs would be identified later in this research project, we prioritized a more iterative and incremental development approach—building and testing small components of the data mart before extending our learnings to larger components (22). Components of the SDLC that we employed were the following: 1) conceptualization, 2) requirements analysis, 3) iterative development, and 4) application release. Our approach to each of these steps is outlined in Table 1.

Project Setting

The setting for this project is the University of Washington Medical Center (UWMC). Through funding received from the Health Resources Services Administration (HRSA) we are implementing IP team-based strategies in HF management for inpatient care at UWMC. The goal of this grant is to improve team functioning, processes of care (e.g., rounds), and HF patient health outcomes. As a first phase in measuring whether introduction of team strategies could affect care processes and ultimately contribute to improved health outcomes, we built an HF 'data mart' containing near real-

Table 1. Software Development Lifecycle (SDLC) steps and HF data mart approach.

SDLC Steps and Definitions	Approach/Activities
Conceptualization: Define the purpose and scope of the HF data mart	Monthly team meetings (starting in October 2014) comprised of interdisciplinary clinical, research and data experts to define the scope and specific questions the data mart would answer; identified existing resources.
Requirements Analysis: Enumerate the different features and functions needed in the data mart	Established common location for sharing information; gained access to institutional DW; obtained secure place to store de-identified HF data mart tables; expanded HF data mart working group.
Iterative Development: Review and refine data extracts	Data Mart: Iterative expansion of HF report; data cleaning and exploration in SQL and STATA; refinement of data dictionary. Team: Subset of working group started meeting weekly in 2015 to coordinate efforts and communication among those working within data mart.
Application Release: Provide access to the data mart to various stakeholders and integrate into work flows	Staged application release: Arm 1: Grant effectiveness (analyzing impact of implementing team-based care strategies on HF patient outcomes); Arm 2: Operations (partnership with QI office, HF nurse educator and HF clinicians); Arm 3: Research (model building, care quality questions, developing protocols and guidelines for collaboration).

time longitudinal inpatient data. In the second phase of this project we will leverage our HF data mart to calculate near real-time predictive analytics and integrate them into the EHR to improve clinical decision making with the goal of yielding improvements in patient outcomes.

Results

Step I: Conceptualization

The purpose and scope for the HF data mart were initially driven by the need to evaluate HF patient outcomes resulting from implementation of IP team-based care strategies. The setting for this work was UWMC, a tertiary/quaternary care academic medical center providing care for inpatients with the full spectrum of HF in the western U.S.

To assess whether HF care processes and patient outcomes improved following implementation of team-based care strategies (e.g., introduction of structured IP bedside rounding, TeamSTEPPS training, and leadership workshops), we required clinical data. To that end, we convened a "data mart working group" in October 2014 to help determine both the resources that already existed, as well as to better understand the broader interest and need for a focused inpatient HF data mart. The data mart "working group" was comprised of an interdisciplinary team that included both academic grant team members and practice-based clinical partners with expertise in informatics, patient safety, QI, health services research, and HF management. The working group identified a concurrent need among clinician partners for an HF data mart to meet multiple investigational, as well as clinical decision-making and QI purposes. In addition, the working group determined that an existing institution-wide DW and HF focused report with underlying queries was already in existence. These entities were important antecedent resources but it was apparent that HF data were not immediately actionable or usable in these existing states. We needed to extract, transform, and load data into a more focused HF data mart that was broader in scope than the daily HF report and narrower than the institution-wide DW.

Step II: Requirements analysis

After clarifying the scope and purpose of the project, the HF data mart working group next conducted a requirements analysis to outline the different features and functions needed in the data mart. Four major categories of requirements emerged in this step: 1) identify a common location for sharing information about HF data mart contents and activities, 2) gain access to an institutional DW, 3) obtain a secure place to store de-identified HF data mart tables, and 4) expand the interdisciplinary working group to develop, utilize, and sustain the HF data mart. These categories are described in more detail below.

1. Establish common location for sharing information about HF data mart contents and activities

Our initial documentation efforts used MS Excel (25) to outline variables included in the initial daily HF report as well as to track the variables we were exploring. In late June, 2015 we migrated the data to a cloud-hosted solution (Google Documents), created a formal data dictionary and documented the status of tables within the HF data mart, as well as analytical datasets exported 'downstream'. These efforts increased ease of access and collaboration for the HF data mart working group who reside within different schools and departments within the institution. A shared folder was used for tracking team priorities and decisions. For each table in the HF data mart, we created a corollary table in the data dictionary where we described metadata about the variables such as a plain language description, technical notes, the data type, and examples. A spreadsheet was created to document a version history for each analytic table that was exported, along with any changes to that table, as

iterations often involved adding new variables to the tables or expanding data ranges. No documents in the shared cloud-based folder included patient-sensitive data.

2. Access to institutional DW

As described in the conceptualization step, preliminary investigations found that the healthcare system had a large institutional DW. Since 2010, the DW team has been working to copy data from disparate operational clinical systems into an institution-wide DW. The DW contains tens of terabytes of source system data, and billions of rows, which are replicated to multiple report servers. Several gigabytes of data and millions of additional rows are added daily (26).

Analysts can query this rich source of consolidated data without affecting the performance of the clinical systems. As seen in Figure 2, depending on the interfaces to the clinical systems, some data is updated in real time, while other data may have a lag or be available on-demand.

Figure 2 reflects feeds in 2015, used with permission of Michael Kuffel (26). Credentialed end-users of these data (clinicians and researchers) who have the necessary skills to connect and query an SQL Server Database (27), can connect to the DW and either extract data into a secure environment for analysis or build secure web-based reports. People meeting these criteria are referred to as 'Affiliate Developers'.

After one member of our HF data mart team became credentialed as an "Affiliate Developer" (November 2014) we applied to the health system to create our own "Collection". Collection is the term the DW team uses to describe a named place within the larger DW where a group can store their own tables and queries saved. This collection of new tables comprises our HF data mart. Co-author Dardas became a 'Collection Owner' in March 2016.

3. Secure place to store de-identified HF data mart tables

Once our "Collection" was established in the DW, we needed a secure place to store analytic files exported from our HF data mart to support data exploration, cleaning and analysis, as well as model
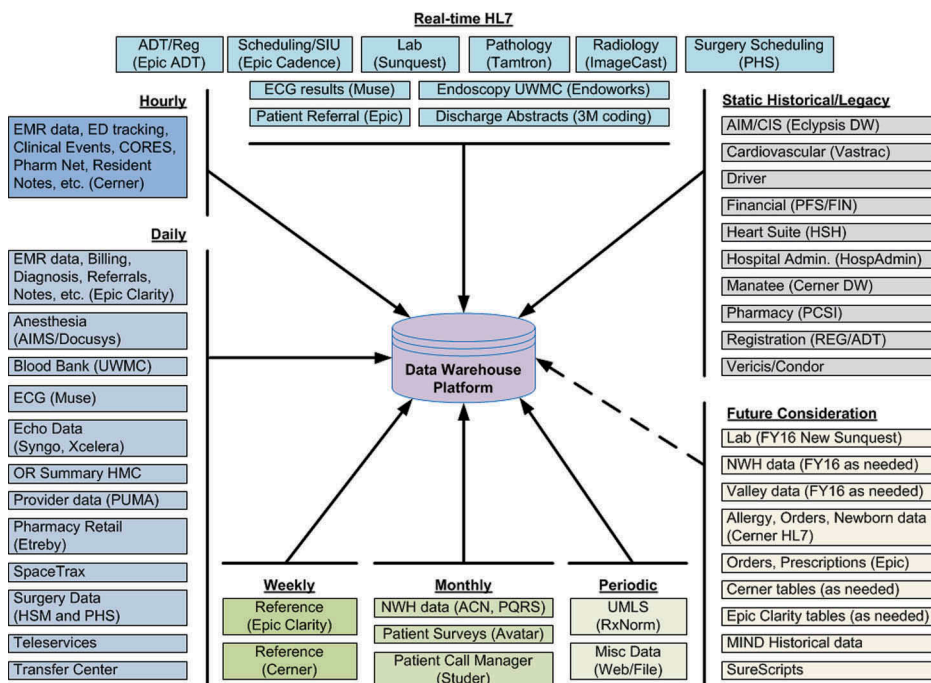

Figure 2. Data warehouse feeds.

building. In collaboration with the institution's Cardiology IT team, the grant HF data mart working group were granted storage space on a HIPAA approved network drive. Exports from our HF data mart collection in the DW are saved to this drive with all data transmission done securely and in compliance with institutional data encryption standards. Using HIPAA safe harbor provisions as a model, before any data sets are exported to this drive, identifiers are removed or obfuscated and all dates are randomly shifted within three days before or after the actual date (28). This date-shifting was implemented at the patient level so that all timestamps across different visits are shifted in unison, making calculations of time-based data, like length of stay and readmission rates possible. We also removed zip code data when counts fell below a certain threshold to protect patient privacy (28).

4. *Expansion of the interdisciplinary HF data mart working group to develop, utilize and sustain the HF data mart*

As described earlier, the HF data mart working group was convened in October 2014. Initial meetings with the grant PI and other clinical and academic stakeholders recurred monthly and focused on clarification of data mart purpose, scope, stakeholders, and resources. We expanded the working group by establishing a partnership with the institution's QI office and invited a representative of that group to join the HF data mart working group. We also identified a need to bring on a data analyst with skills in STATA to assist with data cleaning, exploration, and analysis. To ensure working group clarity and agreement on approaches to handling data, as well as dissemination of results, we also developed data use and team writing agreements. Working group members who needed access to the HF data mart were required to sign the data use agreement; a key element of which was that it prohibited any migration of data off of the secure Cardiology drive.

Step III. Iterative development

Iterative development—the cycle of initial data extracts, review, and refinement—occurred at both the data mart as well as the HF working group level.

As described in earlier steps, we initially identified both the broad institutional DW as well as a narrower existing daily HF report as resources for this project. In order to iteratively develop and expand the HF data mart to meet diverse working group member needs, we decided to utilize and expand the existing HF report.

In 2009, the institution's QI offices developed an HF report for QI to support clinical operations. This report is generated daily to identify HF inpatients across the system at high risk for readmission. Members of the data mart working group met with the HF nurse educator to better understand how she utilized the report, as well as to identify potential areas of improvement that could be introduced/updated using the refined HF data mart.

We also worked with the original analyst (the QI office representative who joined the data mart working group) to repurpose and extend the HF report code to form the basis for the HF data mart. The code worked by extracting data from disparate tables (e.g., patient, visit, labs, medications, imaging, billing) in the institution-wide DW, cleaning and transforming the data, and then appending it to existing tables in our HF data mart. We started with the most recent six months of inpatient visits at one hospital, and then slowly expanded to include all patient visits from two hospitals within the healthcare system. We then exported de-identified versions of these tables 'downstream' to our secure drive as analytic datasets.

Following initial exports, we found that patient data was usually not cleaned or transformed into analyzable formats when pulled from the clinical systems and loaded into the DW. This necessitated adding code to transform the data so that it could be analyzed within statistical software packages. For example, many numerical values are stored as text characters in the

warehouse and we had to develop code to convert these values back to numerical and then quality-test the conversions. When testing approaches to data exploration and cleaning, we found that it was often easier to clean and transform data 'downstream' in STATA as opposed to using SQL in the data mart environment. As more and more data cleaning and exploration occurred 'downstream' in STATA we also established a work process to flag variables so that if they later were determined to be useful 'upstream' (e.g., demonstrated to be significant predictors in a subsequent model), the same process used to clean the data in STATA could be replicated using SQL in the DW. This type of documentation will facilitate hard coding in the DW of data calculations (e.g., risk scores and/or trends) in the HF data mart phase II to enhance availability of real-time decision support tools for clinical care team members. Additionally, we created a table in the dictionary to track all of the exported analytic datasets. For example, we used a 'table overview' to see that on 6/2/2016 we exported version 2.00 of the 'Visits' table and that it contained 206,107 rows of unique visits as well as a number of new variables. Through our iterative development, we have exported over 158 million rows of research data.

In a later iteration of the data dictionary, we also developed and piloted a 'research grade rating' to each variable, scoring it on a scale of 1-4 with a score of "1" indicating there were no known quality issues related to the variable and "4" indicating that there were questions relating to a variable's quality including reliability and/or validity (see Figure 3: Data Dictionary Screenshot). Definitions of these ratings will be refined and formalized as data mart functionality and use increase.

To support these activities and accelerate HF data mart efforts, in 2015, a subset of the HF data mart working group added weekly web-based working meetings to coordinate efforts and communications among the smaller group working directly with the HF data mart. Members of the smaller working group included individuals with expertise in SQL, STATA data analytics, HF, as well as familiarity with the HRSA grant activities and objectives. Increasingly, elements of the HF data mart are becoming functional and able to meet the needs of diverse stakeholders. In the next section, release of aspects of the data mart are described.

Step IV. Application release

As of December 2017, our HF data mart has over 175,200 distinct patients stored in our 'Patients' table; these include both patients with HF and non-HF. Patients without HF were included so that comparative data would be available. Patient characteristics are described in the table using 22 demographic variables drawn from either stable characteristics (e.g., sex) or from

Field Name	Description	Technical Notes	Example	Research grade
AccountStatus	Account Status	Account Status	Discharged	1
AdmitDtTmRC	Admission Date/Time	Date and time - randomly time shifted in synch with Admit Date Time	4/10/2014 17:14:00	2
AdmitSourceDesc	Admission Source	Admission Source	Clinic or Physician's office	2
DemoAge	Age at Admission	Age in years at Admission - randomly time shifted	68	1
DemoDeathDtTmRC	Death Date/Time	Date time for when they died - could have been after hospitalisation. - randomly time shifted	4/11/2014 11:09:00	2

Figure 3. Data dictionary excerpt of visit table.

demographics collected during their latest visit date (e.g., work status). Every patient in this table has one or more rows in the 'Visits' table which currently contains 290,380 rows corresponding to unique visits. This table has 107 variables; ranging from the chief complaint on admission to calculated variables like the weight change between admission and discharge. We also have numerous related 'time series' tables which detail lab values, medication administrations, procedures, and clinical findings like ejection fraction results and blood pressures. Table 2 provides a limited example of variables present in both the 'patients' and 'visits' tables in the data mart and presents a summary of patient characteristics for those with a diagnosis of HF who had an inpatient stay at UWMC between 2010-2016.

Currently, we are following a staged application release, meeting the needs of various stakeholders loosely categorized into three arms with different types of anticipated use:

Arm 1: *HRSA Grant Evaluation*: Through the development of this HF data mart we have been able to assemble patient characteristics and outcomes data needed to evaluate effectiveness of the implementation of team-based care strategies on HF patient outcomes. Since late 2015, the HF data mart has been able to provide de-identified data characterizing the patient population to meet grant-reporting requirements. Increasingly, outcomes of interest are also available in analyzable form including: readmission, mortality, length of stay, as well as data to facilitate analysis of practice variation between and among HF care teams.

Arm 2: *Operations*: In the operations arm we have collaborated with the medical center's QI group to share code and analytic findings that will support updating of existing reports (e.g., the narrower HF report) as well as enhanced capability to flag patients with unidentified and/or undiagnosed HF and those at high risk for readmission and death. For example, we have continued to dialogue with the HF nurse educator to understand the limitations of the existing report from her perspective. She has provided input about what would be helpful to revise or add (e.g., ejection fraction results) that would assist her to more efficiently and effectively identify high risk HF patients in need of targeted education prior to discharge on a daily basis. In this operations arm, we are also working to facilitate development of unit/team specific dashboards that will help HF clinical teams to track clinical pathways that were initiated as well as relevant QI outcomes. We also expect to incorporate risk stratification scores from our models into the dashboards and the daily HF report in phase II of the data mart project.

Table 2. Demographic characteristics of patients with heart failure in data mart (years 2010-2016) (total visits: 26,820; unique patients: 15,131)*

Patient Visit Demographics	n	%
Age (yrs) (mean, sd)	62.32	16.47
Sex (# Male)	16,341	60.93%
Race (# White)	18,696	69.85%
Marital Status (# Married)	11,019	41.37%
Insurance (# Medicare)	13,963	52.06%
Comorbid Conditions (n, %)		
Diabetes	9,685	36.18%
COPD	5,388	20.13%
Depression	5,293	19.77%
CVA	1,230	4.59%
Admission Characteristics		
Length of Inpatient Stay (days) (mean, sd)	9.51	14.28
30-Day Readmissions	5,801	21.63%
Discharge Status		
Home (self)	18,122	67.69%
Skilled Nursing Facility	4,373	16.33%
In Hospital Death	1,943	7.26%
Other	2,334	8.7%

*total "n" ranges from 26,633 to 26,820 for variables presented above based on missing or incomplete data..

Arm 3: *Research*: In this Institutional Review Board (IRB) approved arm, we are focusing on model building and other care quality questions using de-identified datasets exported from our HF data mart. These models may then be integrated into the operations arm if they prove useful in supporting real-time clinical decision-making and/or population health management. We recently received approval to link a retrospective statewide hospital discharge dataset, the Comprehensive Hospital Abstract Reporting System (CHARS) (29). CHARS data contains information about each individual patient's hospital stay. Linking to this information will give us better information about HF patient readmissions to other hospitals as well as non-hospital mortality. We have also developed protocols and guidelines for collaborators to use who are interested in working with our datasets.

Discussion

In order to provide safe, high quality, and effective healthcare, it is imperative that healthcare systems maximize the use of data captured. We described our efforts to create a data mart to support HF management at UWMC to support evaluation of training, QI, and patient outcomes. Other healthcare systems have recognized the need for real-time data marts previously. Intermountain Healthcare has described their experience with developing a DW and decision support tool for management and clinical decision-making, including for identification of patients at high risk for venous thromboembolism and infection control (30). Other healthcare institutions including Mayo Clinic (31) and Partners HealthCare System in Boston (21) have previously developed their own DW in attempts to develop shared, standards-based infrastructure for clinical care and research purposes. The success of our current efforts at UWMC have hinged on careful data governance through a multidisciplinary team that included developers familiar with the complexities of our healthcare information network, with content experts in health services and biostatistics, working closely with HF cardiologists and nurse practitioners since the initiation of our project. We intentionally and prospectively identified key variables, studies, and endpoints that need to be accurately and appropriately extracted to maximize data integrity and meet our funding and project focus. Through this process, we have assured consistent model and content development and the creation of a data mart that is readily applied to clinical use in our targeted HF population. Initial outcomes from each of the three arms are promising.

Key Findings from Phase 1

We believe a number of lessons learned through this process are generalizable to other researchers interested in similar pursuits (e.g. with other acute diseases). Foremost, we found it feasible to build an HF data mart and to meet the information needs of diverse stakeholders. However, it is also a complicated process that requires a team with multiple types of expertise (e.g., clinical knowledge, coding in SQL, data management, research). Building of a dedicated HF data mart also requires a high level of investment in data quality/testing, as initial crude processes are refined and data quality issues become known. Despite these challenges, there is a clear desire and need for this kind of data, as well as a potential for sharing structures, processes, and outcomes of actual data.

One challenge we have faced is maintaining a clear direction with our research questions. There are many variables to work with and stakeholders with disparate needs and as a result, it is easy to explore tangents and to lose sight of primary questions. Nonetheless, this is part of the learning process and we see this as an important step in supporting an academic-practice partnership.

Another challenge we have faced is locating new variables in the DW and understanding their operational definitions. Work is underway to make things easier; the DW team has also been building a semantic layer on top of the databases that provides data in more user-friendly object models. However, these object models remain in development by the DW team, as they become

available to us we will refine our code to pull data from these object models instead of the raw tables, thus requiring less data cleaning transformations on our side.

Much of the classification we are currently analyzing is based on how diagnoses were coded after patients were discharged. We need to determine better ways to identify patients with HF before they are discharged. Currently, admitting diagnosis fields are not well structured making it challenging to identify patients who may have HF, but are not admitted to an HF service (for example, a patient who is admitted for a fracture but has a history of HF that may be exacerbated during treatment of the fracture).

Our HF data mart is similar to a patient registry as it contains a focused list of patients with HF and related visit information. However, at this point we are unable to capture events that occur outside of this health system. For example, if a patient seeks care at this institution (UWMC is a referral center for HF care for a 5-state region) and then returns to their home town, we will not be able to access or incorporate data that results from visits elsewhere. Many national and international HF registries have been built to better understand and address the challenges faced by a growing HF population in the U.S. and worldwide (33, 34). In the U.S. these datasets have included models such as, ADHERE (35), OPTIMIZE-HF (36), and GWTG-HF (37).

Despite the expansive nature of these registries, there remain important gaps that cannot be addressed due to methods of implementation and data collection. The majority of these registries have retrospective, nonconsecutive enrollment of patients, which may result in differential recruitment of HF patients and may not be representative of the entire HF population. Data collection is typically manually performed, resulting in potential transcription errors and/or missing data. Longitudinal data may not always be collected and post-discharge outcomes may be missing, but possibly obtained by linking to external national reporting databases.

Our HF data mart is bridging many of these gaps and making it possible to address questions that are problematic to answer using large-scale registries. Our HF data mart captures all HF patients consecutively that enter our healthcare system and data collection is automated via iterative extraction from the DW. In the future, we anticipate that longitudinal data collected for patients that remain within our healthcare system post-discharge will be added to the database, and we will have real-time identification and risk estimates for our HF cohort.

In Phase II, we anticipate being able to explore and build more robust models, including ones that look at the slopes of variables like weight change, rather than variables that are essentially static or cross-sectional (snapshots in time). These models will help us to refine techniques for prospective identification of HF patients as well as identify those at higher risk for complications or readmissions through development of algorithms.

We also anticipate interfacing with other data within the institution to expand data capture — including outpatient data systems and building interfaces for manually adding, or batch importing data (e.g., patient reported outcomes data) to align and expand the capacity of our HF data mart to better meet the needs of all stakeholder groups above. Addition of this type of information is consistent with the national direction in which registries and other data marts are being used to support decisions in real-time and assist in the management of population health. The ability to expand data capture also enhances the potential to expand the data mart to other acute diagnoses and patient populations.

Limitations

As with any study or project, there are limitations of this data mart. First, as is well documented, medical records' data are often characterized by missing or inaccurate information (38). Additionally, a limitation for this project has been the identification of HF patients in real time as diagnosis codes are not assigned until after discharge. A third limitation is that we are unable to capture events that occur outside of our healthcare system.

Conclusions

HF is a significant public health problem that is increasing. Secondary and tertiary prevention can be supported by tapping into large clinical DW. This requires specialized tools and skills to extract the data and transform it into actionable information that can then be used by QI and research teams, as well as clinicians. Our first phase of data mart work also sets the stage for using predictive analytics to help identify efficient care patterns, practice variations, and for risk stratification. We see this project as an initial step in these directions and look forward to sharing our steps and learnings on completion of the second phase of our project. This is a replicable model that can be used in other settings and for other diagnoses.

Funding

References

1. Berwick DM, Nolan TW, Whittington J. The triple aim: Care, health, and cost. Health Affairs (Millwood) 2008;27(3):759–69. doi: 10.1377/hlthaff.27.3.759
2. Institute of Medicine of the National Academies. Best care at lower cost: The path of continuously learning health care in America; 2012. Available from: http://tiny.cc/dsawky
3. Terry AL, Cejic S, Ryan BL, Shadd JD, Stewart M, Fortin M, Thind A. You and your EMR: The research perspective Part 4. Optimizing EMRs in primary health care practice and research. Canadian Family Physician 2012;58(6), 705–706. Available from: http://www.cfp.ca/content/58/6/705
4. Turley CB, Obeid J, Larsen R, Fryar KM, Lenert L, Bjorn A., . . . Sanderson I. Leveraging a statewide clinical data warehouse to expand boundaries of the learning health system. EGEMS (Washington, DC) 2016;4(1), 1245. doi. org/10.13063/2327-9214.1245
5. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, . . . Turner MB. Heart disease and stroke statistics-2016 update: A report from the American Heart Association. Circulation 2016;133:e38–360.
6. Desai AS, Stevenson LW. Rehospitalization for heart failure: predict or prevent? Circulation 2012;126(4), 501–506. https://doi.org/10.1161/CIRCULATIONAHA.112.125435
7. Cowie, M. R., Anker, S. D., Cleland, J. G. F., Felker, G. M., Filippatos, G., Jaarsma, T., . . . López-Sendón,J. (2014). Improving care for patients with acute heart failure: before, during and after hospitalization. ESC Heart Failure, 1(2), 110–145. https://doi.org/10.1002/ehf2.12021
8. Von der Heidt, A., Ammenwerth, E., Bauer, K., Fetz, B., Fluckinger, T., Gassner, A., . . . Poelzl, G. (2014). HerzMobil Tirol network: rationale for and design of a collaborative heart failure disease management program in Austria. Wiener Klinische Wochenschrift, 126(21–22), 734–741. https://doi.org/10.1007/s00508-014-0665-7
9. Committee, H. F. P. W., Yancy, C. W., Januzzi, J. L., Allen, L. A., Butler, J., Davis, L. L., . . . Wasserman, A. (2017). 2017 ACC Expert Consensus Decision Pathway for Optimization of Heart Failure Treatment: Answers to 10 Pivotal Issues About Heart Failure With Reduced Ejection Fraction. Journal of the American College of Cardiology, 24465. https://doi.org/10.1016/j.jacc.2017.11.025
10. Tait, G. R., Bates, J., LaDonna, K. A., Schulz, V. N., Strachan, P. H., McDougall, A., & Lingard, L. (2015). Adaptive practices in heart failure care teams: implications for patient-centered care in the context of complexity. Journal of Multidisciplinary Healthcare, 8, 365–376. https://doi.org/10.2147/JMDH.S85817
11. LaDonna, K. A., Bates, J., Tait, G. R., McDougall, A., Schulz, V., Lingard, L., & Heart Failure/Palliative Care Teamwork Research Group. (2017). "Who is on your health-care team?" Asking individuals with heart failure about care team membership and roles. Health Expectations: An International Journal of Public Participation in Health Care and Health Policy, 20(2), 198–210. https://doi.org/10.1111/hex.12447
12. Browne, S., Macdonald, S., May, C. R., Macleod, U., & Mair, F. S. (2014). Patient, carer and professional perspectives on barriers and facilitators to quality care in advanced heart failure. PLoS One, 9(3), e93288. https://doi.org/10.1371/journal.pone.0093288

13. Simmonds, R., Glogowska, M., McLachlan, S., Cramer, H., Sanders, T., Johnson, R., … Purdy, S. (2015). Unplanned admissions and the organisational management of heart failure: a multicentre ethnographic, qualitative study. *BMJ Open*, 5(10), e007522. https://doi.org/10.1136/bmjopen-2014-007522

14. Organizations, J.C.o.A.o.H., *Improving America's Hospitals: The Joint Commission's Annual Report on Quality and Safety 2010*. 2010: Joint Commission.

15. Institute of Medicine (US) Committee on Quality of Health Care in America. *To err is human: Building a safer health system*. (L. T. Kohn, J. M. Corrigan, & M. S. Donaldson, Eds.). (2000). Washington (DC): National Academies Press (US). Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK225182/.

16. James, J.T., *A new, evidence-based estimate of patient harms associated with hospital care*. Journal of Patient Safety, 2013. 9(3): p. 122–128.

17. Barceló, A., et al., *Using collaborative learning to improve diabetes care and outcomes: the VIDA project*. Primary care diabetes, 2010. 4(3): p. 145–153.

18. Narayan M, Jones J, Portalupi LB, McIlvennan CK, Matlock DD, Allen LA. Patient perspectives on communication of individualized survival estimates in heart failure. Journal of Cardiac Failure 2017;23(4), 272–277. https://doi.org/10.1016/j.cardfail.2016.12.008

19. Inmon W. Building the data warehouse. John Wiley & Sons, Inc. New York, NY, USA; 1992.

20. Inmon W. Data mart does not equal data warehouse. DMReview.com; 07/18/00. Available from: https://goo.gl/XFSD6j

21. Gliklich RE, Dreyer NA, Leavy MB. Executive Summary. Agency for Healthcare Research and Quality (US); 2014. Available from: http://tiny.cc/i9bwky

22. Davis AM, Bersoff EH, Comer ER. A strategy for comparing alternative software development life cycle models. IEEE Transactions on Software Engineering 1988;14(10), 1453–1461. https://doi.org/10.1109/32.6190 http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6190&tag=1

23. Royce WW. Managing the development of large software systems. Proceedings of IEEE WESCON 1970;26 (August): 1–9. Available from: http://tiny.cc/r1awky

24. Larman and Basili. Iterative and incremental development: A brief history. 2003. Available from: http://tiny.cc/52awky

25. Microsoft and GE healthcare complete joint venture agreement. 2012, June 6. Available from: http://tiny.cc/l4awky

26. Kuffel, M. Written communication (2017, February 15).

27. SQL Server 2012, product documentation and technical articles. 2012. Available from: https://technet.microsoft.com/en-us/library/dn197878(v=sql.10).aspx

28. Department of Health and Human Services [Internet] (n.d.) Available from: https://goo.gl/sX0tvX. Retrieved on 04/13/17

29. Washington State Department of Health [Internet] (n.d.) Available from: https://goo.gl/NTnvn3

30. Evans RS, Lloyd JF, Pierce LA. Clinical use of an enterprise data warehouse. AMIA Annual Symposium Proceedings 2012;189-198.

31. Chute CG, Beck SA, Fisk TB, Mohr DN. The enterprise data trust at Mayo Clinic: A semantically integrated warehouse of biomedical data. Journal of the American Medical Informatics Association 2010;17:131–135.

32. Murphy SN, Barnett GO, Chueh HC. Visual query tool for finding patient cohorts from a clinical data warehouse of the partners healthcare system. AMIA Annual Symposium Proceedings 2000;1174.

33. Ambrosy AP, Fonarow GC, Butler J, Chioncel O, Greene SJ, Vaduganathan M, Nodari S, … Gheorghiade M. The global health and economic burden of hospitalizations for heart failure: Lessons learned from hospitalized heart failure registries. Journal of the American College of Cardiology 2014;63:1123–1133.

34. Maggioni AP, Dahlstrom U, Filippatos G, Chioncel O, Leiro MC, Drozdz J, Fruhwald F, … Tavazzi L. Eurobservational research programme: The heart failure pilot survey (esc-hf pilot). European Journal of Heart Failure 2010;12:1076–1084.

35. Adams KF, Jr., Fonarow GC, Emerman CL, LeJemtel TH, Costanzo MR, Abraham WT, Berkowitz RL, … Horton DP. Characteristics and outcomes of patients hospitalized for heart failure in the united states: Rationale, design, and preliminary observations from the first 100,000 cases in the acute decompensated heart failure national registry (adhere). American Heart Journal 2005;149:209–216.

36. Fonarow GC, Abraham WT, Albert NM, Gattis Stough W, Gheorghiade M, Greenberg BH, O'Connor CM, … Young JB. Influence of a performance-improvement initiative on quality of care for patients hospitalized with heart failure: Results of the organized program to initiate lifesaving treatment in hospitalized patients with heart failure (optimize-hf). Archives of Internal Medicine 2007;167:1493–1502.

37. Smaha LA. The American Heart Association gets with the guidelines program. American Heart Journal 2004;148:S46–48.

38. Bowman, S. Impact of electronic health record systems on information integrity: Quality and safety implications. Perspectives in Health Information Management 2013;10(Fall). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3797550/