Visualizing collaborative electronic health record usage for hospitalized patients with heart failure

Nicholas D Soulakis[1], Matthew B Carson[1,2], Young Ji Lee[1], Daniel H Schneider[3], Connor T Skeehan[1], Denise M Scholtens[4]

AMIA INFORMATICS PROFESSIONALS. LEADING THE WAY. **OXFORD** UNIVERSITY PRESS

ABSTRACT

Objective To visualize and describe collaborative electronic health record (EHR) usage for hospitalized patients with heart failure.

Materials and methods We identified records of patients with heart failure and all associated healthcare provider record usage through queries of the Northwestern Medicine Enterprise Data Warehouse. We constructed a network by equating access and updates of a patient's EHR to a provider-patient interaction. We then considered shared patient record access as the basis for a second network that we termed *the provider collaboration network*. We calculated network statistics, the modularity of provider interactions, and provider cliques.

Results We identified 548 patient records accessed by 5113 healthcare providers in 2012. The provider collaboration network had 1504 nodes and 83 998 edges. We identified 7 major provider collaboration modules. Average clique size was 87.9 providers. We used a graph database to demonstrate an ad hoc query of our provider-patient network.

Discussion Our analysis suggests a large number of healthcare providers across a wide variety of professions access records of patients with heart failure during their hospital stay. This shared record access tends to take place not only in a pairwise manner but also among large groups of providers.

Conclusion EHRs encode valuable interactions, implicitly or explicitly, between patients and providers. Network analysis provided strong evidence of multidisciplinary record access of patients with heart failure across teams of 100+ providers. Further investigation may lead to clearer understanding of how record access information can be used to strategically guide care coordination for patients hospitalized for heart failure.

RESEARCH AND APPLICATIONS

BACKGROUND AND SIGNIFICANCE

Well-coordinated, collaborative, multidisciplinary care improves patient outcomes and decreases medical costs.[1–3] Federal agencies including the Centers for Medicare and Medicaid Services and the Agency for Healthcare Research and Quality (AHRQ) seek to increase coordinated care nationally as a means of improving healthcare quality.[4] However, the ability to define and measure coordinated care using electronic health records (EHRs) remains elusive.[5–7] The AHRQ's *Care Coordination Measures Atlas* suggests that care coordination measures depend on health information technology (HIT) systems that track essential data elements.[8] Identifying the best strategies to leverage HIT systems for ascertaining and monitoring care coordination is essential, particularly in populations requiring a disproportionate share of healthcare spending in the United States, as is the case for patients with heart failure at risk of readmission.[9–15]

Several approaches to measuring coordinated care have been put forth, including process measures for patient education, self-management, nutritional counseling, and intensive follow-up by telephone.[16–20] Predominant approaches measure the effects of these care coordination components through case study,[20] analysis of reimbursement data,[22–27] or patient survey.[28–30] A recent literature review of network analysis in healthcare studies showed that 50 of 52 studies used survey or observation to collect data; one study used physicians' orders from the health information system and another study utilized process logs.[31] Manual data collection approaches are not suitable for ongoing, systematic evaluation[32] and furthermore lack comprehensive ascertainment of realized care coordination at

For numbered affiliations see end of article.

healthcare facilities. A powerful alternative lies in the administrative data found within EHRs, which is useful for managing identity and access. This has been brought to light by the recent Strategic Health IT Advanced Research Projects initiatives funded by the Office of the National Coordinator for HIT.[33–36] Briefly, clinical systems enable capture of EHR access by all users whether this access involves only browsing a record or updating it. This unique data source has been shown to effectively identify out-of-the-ordinary record access by a variety of clinical and administrative actor types as well as use of clinical documentation.[33,37–40]

Recent literature provides strong evidence that thoughtfully designed care coordination with explicitly defined multidisciplinary intervention teams demonstrate positive effects on measures related to hospital readmissions for the chronically ill.[41–43] However, top-tier healthcare providers respond to new challenges and innovate on a daily basis to create novel patterns of care. To date, no study has presented a clear, reproducible method for discovering these spontaneous, contemporaneously assembled, and self-organized models of healthcare.

Recently, graph theory and network analysis have been applied to a variety of large data sets in order to elucidate the nature and mechanisms of relationships in complex systems.[44–48] Of particular interest is the study of collaboration networks, which are graphs representing a social structure based on some type of shared interest. Recent publications have utilized bipartite networks to model common activities among individuals and to describe collaboration networks in diverse settings from Broadway musicals to scientific paper authorship.[49–52] This same method can be applied to healthcare coordination within a hospital setting by creating a provider-patient– directed bipartite network and extrapolating from this a provider collaboration network.

OBJECTIVE

Our goal is to ascertain and monitor care coordination by developing a graph-based analysis method to identify healthcare interactions among populations known to experience high readmission rates. As a first step toward that goal, the objective of this study was to visualize and describe collaborative electronic health record usage for hospitalized patients with heart failure.

MATERIALS AND METHODS

The Northwestern Memorial Hospital (NMH) is an 876-bed teaching institution for the Feinberg School of Medicine at Northwestern University. The institution's Enterprise Data Warehouse (EDW) was used to identify patients aged between 18 and 89 years with a primary billing diagnosis of heart failure using ICD9 codes selected for quality initiatives at the Center for Heart Failure within the Northwestern Medicine Bluhm Cardiovascular Institute. Inpatient and observation encounter types not resulting in a readmission that had a discharge date between January 1, 2012 and January 1, 2013 were included. Patients admitted to the psychiatric wards were omitted based on Illinois state law regarding Sensitive Protected Health Information. For each patient, additional data collection

included length of stay, death status, discharge disposition, and an intensive care unit (ICU) indicator. Providers were determined by identifying those that interacted with heart failure study patients by accessing their clinical content using the Cerner Corporation's PowerChart (Cerner Corporation, Kansas City, MO, USA) electronic medical record system. Healthcare "providers" include all healthcare professionals and nonclinical staff. While our definition does not fit with the strict definition of a healthcare professional, we chose it to maintain the terminology convention used within the Cerner system. We will refer to all NMH personnel as providers for the remainder of the paper. Position, access type (order vs nonorder), and a physician indicator were collected. "Position," also a Cerner convention, refers to the highest level of provider type. We identified employees with the ability to either place an order (ie, perform an order action) or be assigned responsibility for an order by having that order given to them. If a user did not have one of these options, we considered them to have nonorder access. Northwestern University's Institutional Review Board approved the study with a waiver of patients' informed consent.

Pipeline

Figure 1 is an illustration of the pipeline used in this study. Information was initially retrieved from EHRs and transferred into an operational data store within the Northwestern Medicine EDW using extract, transform, and load (ETL) scripts. Microsoft SQL Server Management Studio[53] was then used to query and export the data set for this study from the EDW to comma separated value (.csv) files. These files were subsequently parsed and edited using scripts written in Perl[54] and Python[55] to create input files for Gephi,[56] R,[57] and Neo4j[60] for further analysis. Neo4j graph database queries were written using the native query language, Cypher.

Provider and Patient Attributes

We collected a set of attributes for providers and patients from the EDW. For providers, we extracted employee ID, employee role, and a physician index (1 if provider is a type of physician, 0 if not). For patients, we extracted age, encounter type (all inpatients in this data set), admission and discharge times, primary diagnosis, discharge location, length of stay, medical service (department to which the patient was admitted), discharge disposition (where the patient was discharged to), and an index that noted whether or not the patient had died.

Network Visualization

We created 2 types of networks in this study. The first is a directed bipartite network and represents interactions between providers and patient records (figure 2). The second network is undirected and depicts shared patient record access between providers (figure 3). Visualization for both networks was performed using Gephi.[56] Further description of these networks follows.

Provider-patient Network

We created a directed bipartite graph that depicts shared record access for a single patient with an average number of

Figure 1: An illustration of the pipeline used in this study.
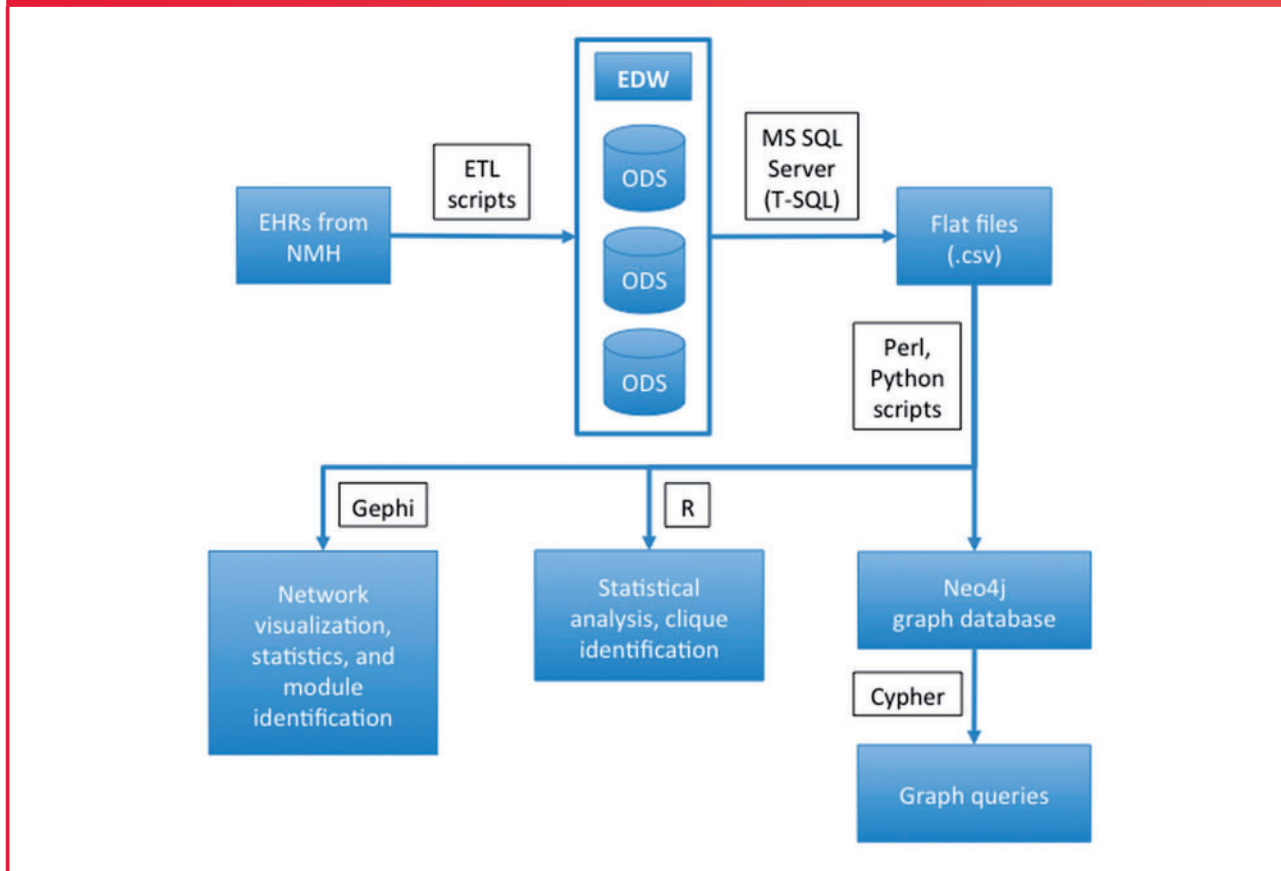
provider interactions (or average team size) and a typical length of stay (figure 2). Source nodes designate a type of provider— "physician," "nurse," "pharmacist," etc. The target node (center) designates a diagnosis of heart failure in a patient who was admitted to NMH in 2012. An edge between a provider and a patient is an indication that the provider accessed the patient record. Nodes are colored by provider position and edge lengths indicate the number of orders (procedures, prescriptions, laboratory tests, or consultations) submitted by the providers. For our purposes, "shared record access" for any given patient was defined as all providers who accessed his or her record. The complete provider-patient network included all providers as the set of source nodes and all patients as the set of target nodes.

Provider Collaboration Network

An undirected provider collaboration network was developed using data recorded in the provider-patient network. Nodes in the provider collaboration network represent providers and edges represent pairwise shared access of at least 10 patient records. We chose to exclude shared access of less than 10 records for two reasons. First, omitting these edges removed potential noise and improved data consistency. Second, it greatly

reduced the number of interactions within the network (from 2 217 570 edges to 83 998 edges) and made our analysis computationally feasible. Edges in the provider collaboration network were weighted by a shared patient index (SPI)

$$SPI = \frac{|PR_A \cap PR_B|}{|PR_A \cup PR_B|},$$

where PR_A is the set of patient records accessed by Provider A, PR_B is the set of records accessed by Provider B, and $0 \leq SPI \leq 1$.

Also known as the Jaccard index,[59] this statistic measures the similarities and differences between the 2 sets. By adding an edge weight, we treated some interactions as having "stronger ties" than others. Specifically, the relationship between 2 providers who had a higher percentage of their total number of patient record accesses in common was stronger than the relationship between 2 providers who had less patient record accesses in common. We believe that this method more accurately represented provider interactions than treating all relationships equally.

Nodes and edges in figure 3A were arranged using a 2-step process. First, we applied the Force Atlas layout.[56] This force-directed algorithm considers edge weights in determining the strength of a relationship between 2 nodes. A stronger

RESEARCH AND APPLICATIONS

Figure 2: (A) An example of a provider-patient network for a single patient. Number of providers = 85; Number of order-placing providers = 15; Other = Administrative staff and Quality Assurance; PCSN = Patient Care Staff Nurse; R/F = NMH Resident/Fellow; PCAS = Patient Care Assistive Staff; Resp. = Respiratory; APC = Advanced Practice Clinician; Hosp. = Hospitalist; PA = Physician Assistant; Rad. = Radiologist; US = Unit Secretary. An edge connects a provider and the patient if the provider accessed the patient's record. Provider nodes in closer proximity to the patient node indicate that the provider placed a higher number of orders for the patient. (B) A network for the same patient including only providers designated as physicians (number of physicians = 24).
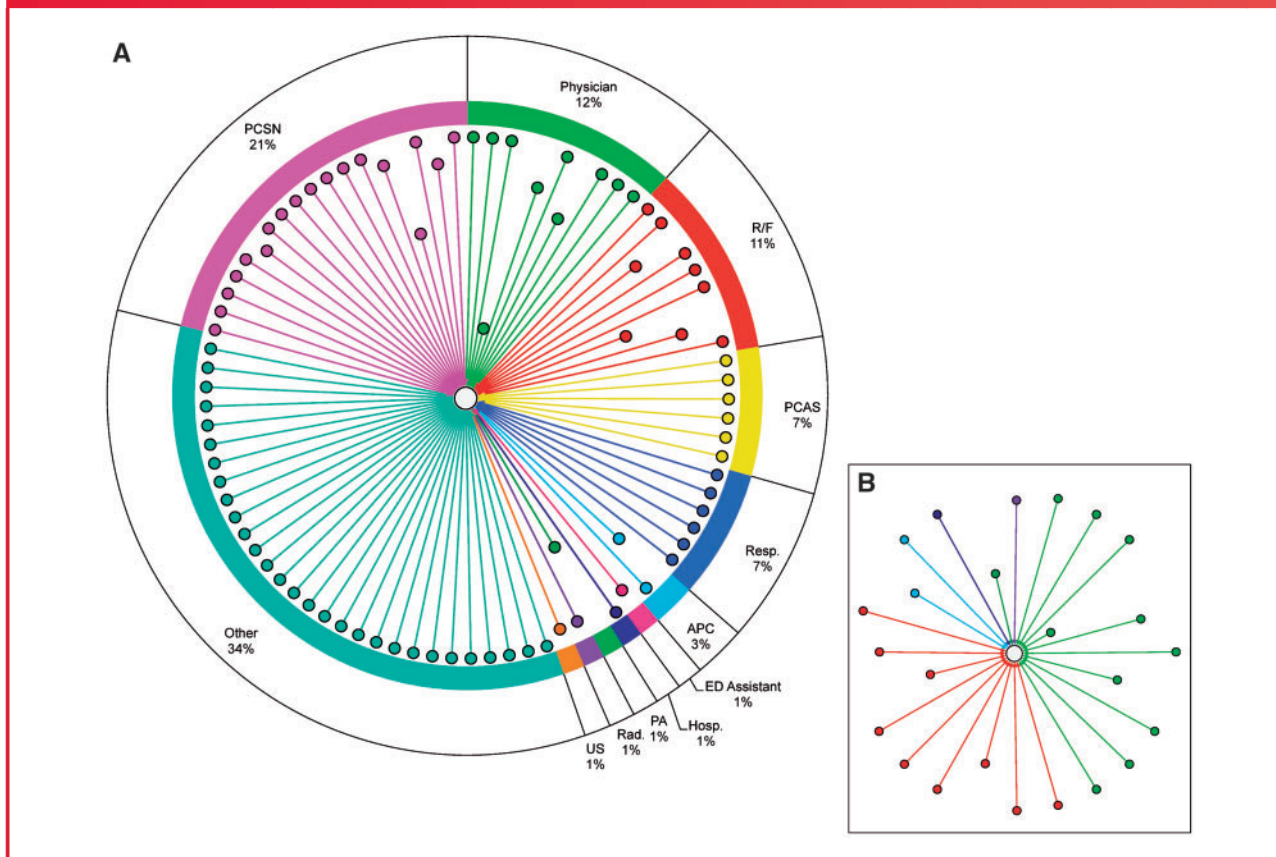
relationship (larger edge weight) between 2 nodes resulted in a shorter edge length between them. Weaker relationships resulted in longer edge lengths. Second, we applied the Fruchterman Reingold layout,[60] another force-directed algorithm implemented in Gephi.[56] This layout did not use edge weights but instead attempted to distribute nodes evenly within the graph area to allow as few overlapping edges as possible and to make all edge lengths equal. The reason for the 2-step process was that the output generated by force-directed algorithms was influenced by the initial layout. By using the Force Atlas algorithm first, we took into account edge weights and were able to group nodes by module (color). The Fruchterman Reingold algorithm then modified the output of the first algorithm by making the graph roughly symmetrical with edges of approximately equal length. In figure 3, edges were assigned the same color as the source node, and edge length was not connected to any attribute.

Module and Clique Identification
Sets of providers with high frequencies of pairwise shared record access were identified using 2 approaches. The first used the heuristic community detection algorithm by Blondel et al[61] to identify sets of providers, or modules, with a higher density of edges within their set than outside the set (figure 3B). The algorithm took into account edge weights and used randomization when identifying modules. We used the suggested resolution value of 1.0. As before, edges were assigned the same color as the source node. Edge width was tied to the number of intermodule interactions (shared record accesses between providers). A thicker edge indicates more interactions between providers in different modules (eg, modules 1 and 2 shared more intermodule provider interactions than 4 and 6; figure 3B).

The second approach adopted a stricter definition for shared record access by identifying all cliques in the data set (figure 4). For our purposes a clique was defined as the largest

Figure 3: Two representations of the patient-sharing network are shown. Node colors identify module membership. Edges are assigned the color of the source node. (A) Network view: nodes are providers and edges between providers indicate ≥ 10 shared patient record accesses. The edge weight was calculated using the shared patient index (see *Materials and Methods* section). (B) Modular view: nodes are groups of providers (modules) and edges between modules indicate patient record sharing between members of 2 modules and are weighted by total number of interactions between members. A total of 7 modules were identified.

RESEARCH AND APPLICATIONS
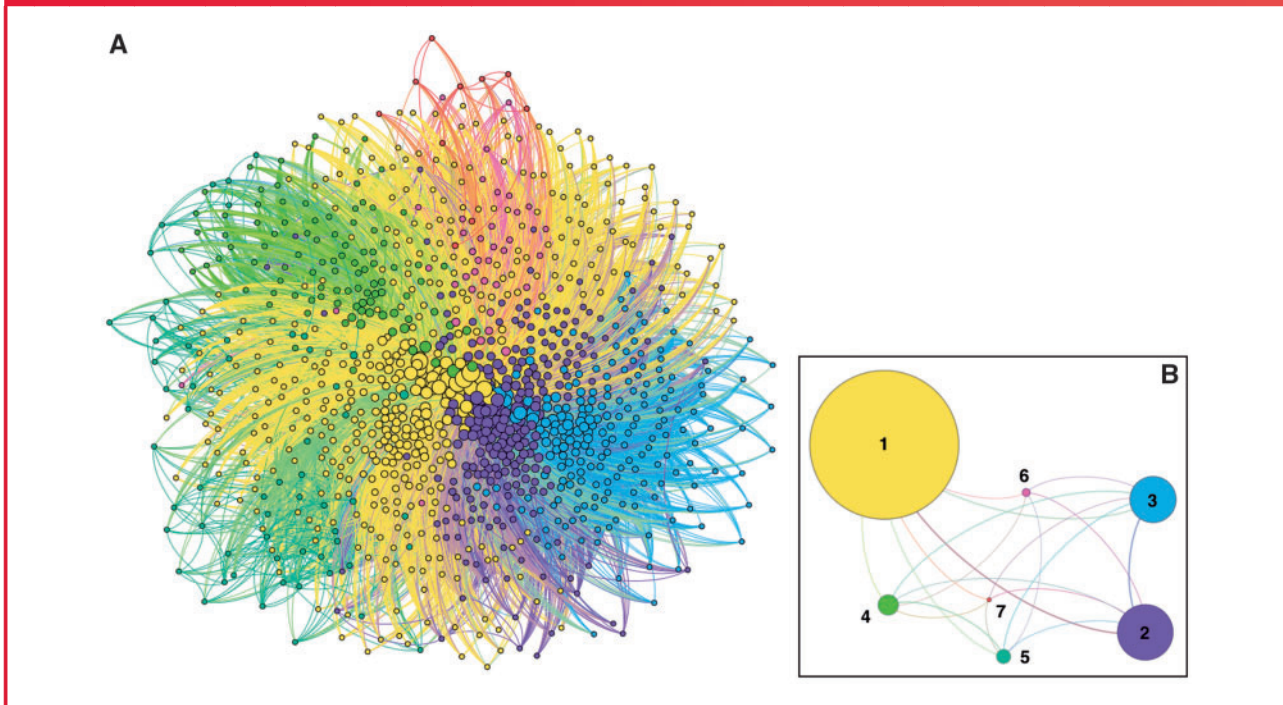
set of providers who shared record access with all other providers in the set. Whereas a single edge identifies shared record access between 2 providers, a clique reveals the extent of shared record access on a more global level. Clique membership and size within the provider collaboration network were identified using the kCliques algorithm,[62] which is part of the Boost Graph Library interface package (RBGL)[63] for the R statistical computing environment.[57]

Querying the Provider-Patient Network

To facilitate the data mining of provider-patient and care collaboration networks, we built a database of patients, providers, and interactions using the Neo4j graph database platform (figure 5). A graph database combines the advantages of a traditional relational database with the high-performance capabilities of graph search methods.[64] The database can be queried using a native query language (Cypher), which is similar to SQL. We included all patient and provider properties mentioned previously in the *Materials and Methods* section. As a guide for conducting ad hoc queries of the graph database, several data mining packages were used to partition the data set based on predefined outcome measures such as mortality, ICU stay, or a length of stay > 7 days (R:rpart,[65] WEKA: J48 decision tree,[66] and Salford Predictive Suite: CART[67,68]).
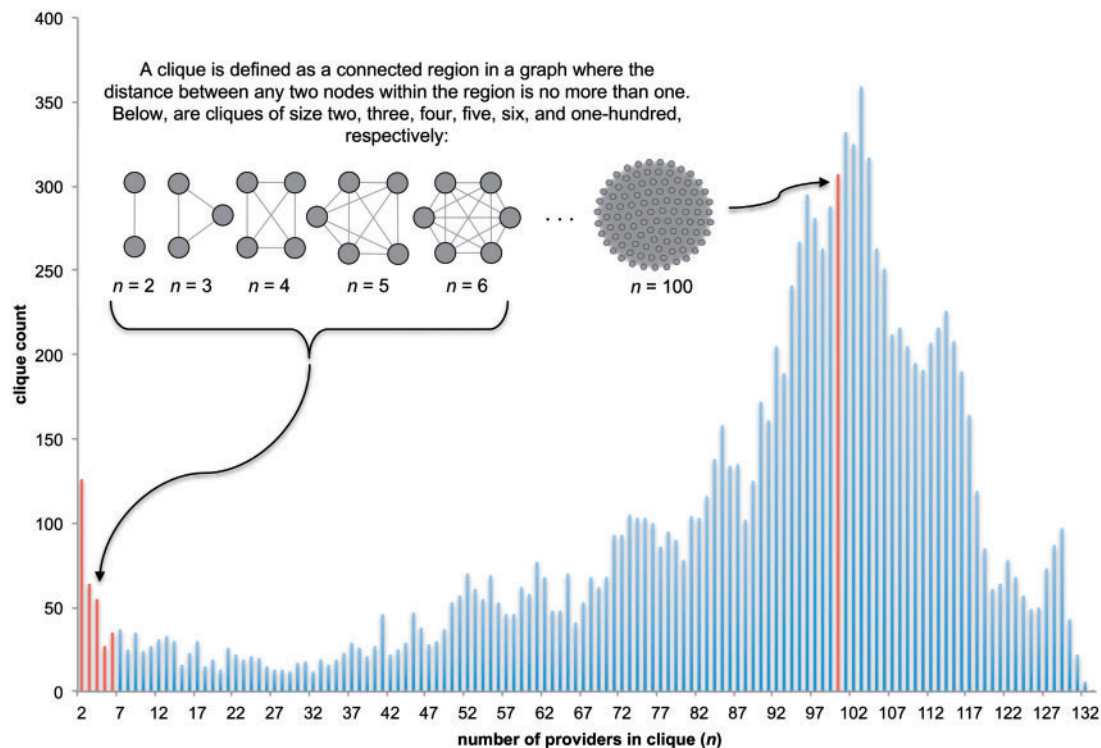
All provider position types were entered as predictors and outcome measures were used as target variables. The results from each tool were compared to identify the best decision rules. Rules of interest were used to form Cypher queries. The Neo4j explorer interface was then used to further examine subnetworks.

RESULTS

The initial data warehouse query provided us with a cohort of 695 patients with distinct heart failure. After checking data quality and excluding readmissions, we identified 548 patients meeting the study criteria. The average patient age was 65 years and the average length of stay was 8 days, with 374 of 548 patients (68%) discharged in ≤ 7 days. Twenty-one of 548 patients (4%) died and 153 of 548 patients (28%) were admitted to the ICU at some point during their hospitalization. We identified 5113 healthcare providers representing 131 unique positions. One thousand four hundred fifty-four of 5113 (28%) providers were physicians.

A summary of shared access statistics for the most highly represented provider positions in the patient-provider network is shown in table 1. *Providers* indicates the number of providers in the position; *Patients* is the number and percentage of patient records a provider of this type accessed; *Median*

Figure 4: The results of clique identification for the provider collaboration network in Figure 3A are shown. "Number of providers in clique (n)," or clique size, is plotted against "clique count," or the number of cliques of size n. Cliques are defined and illustrated in the upper left hand corner. Nodes represent providers and an edge between 2 providers indicates that they accessed at least 10 of the same patient records. The clique identification algorithm did not take the shared patient index edge weight into account. The red bars indicated by the arrows designate the number of cliques of size 2, 3, 4, 5, 6, and 100 identified in the network.


A clique is defined as a connected region in a graph where the distance between any two nodes within the region is no more than one. Below, are cliques of size two, three, four, five, six, and one-hundred, respectively:

n = 2   *n* = 3   *n* = 4   *n* = 5   *n* = 6   *n* = 100

providers/patients (min, max) indicates the median, minimum, and maximum number of providers of this type who accessed a patient record; *Number record accesses* is the total number of provider-patient relationships for a position in the network; and *Median patient record accesses (min, max)* indicates the median, minimum, and maximum of the previous value. The following top 5 most frequently observed provider positions made up 46% (2384 of 5113) of the total positions: *Patient Care Staff Nurse* (n = 940; 18%, 940 of 5113), *NMH Resident/Fellow* (n = 654; 13% 654 of 5113), *Medical Student* (n = 290; 6%, 290 of 5113), *NMH Physician* (n = 262; 5%, 62 of 5113), and *Patient Care Assistive Staff* (n = 238; 5%, 238 of 5513). The position with the largest number of providers was *Patient Care Staff Nurse*, and each interacted with a median of 12 patients. In contrast, the *Respiratory* position had a total of 107 providers in the network, but each provider interacted with a median of 12 patients. We identified a total of 61 026 patient record accesses in our data set. The following 5 provider positions accounted for 52% (32004 of 61026) of these record accesses: *Patient Care Staff Nurse* (n = 9014; 15%, 9014 of 61026), *Respiratory* (n = 8761; 14%, 8761 of 61026), *NMH*

Resident/Fellow (7000; 11%, 7000 of 61026), *Patient Care Assistive Staff* (n = 4662; 8%, 4662 of 61026), and *Rx-Pharmacist* (n = 2567; 4%, 2567 of 61026).

We defined another collaboration type, multidisciplinary patient record sharing, as one in which a provider shares patient records with another provider of a different position type. Table 2 summarizes collaborations across the entire network as well as those collaborations in which each provider shares at least 10 patient record accesses with another provider (described further in figure 2). For the comprehensive collaboration network, the top 5 positions with the highest median number of multidisciplinary collaborations per provider were *Respiratory* (n = 2711), *RX-Pharmacist* (n = 1709.5), *RAD-Nurse* (n = 1067.5), *Patient Care Assistive Staff* (n = 832), and *NMH Resident/Fellow* (n = 721.5). For the collaboration network with ≥ 10 shared patients, the top 5 positions in this respect were *Respiratory* (n = 283), *Advanced Practice Clinician* (n = 204), *RX-Pharmacist* (n = 128), *NMH Anesthesia* (n = 82), and *RAD-Radiologist* (n = 56). Despite a high number of multidisciplinary collaborations per provider in the comprehensive network, the *RAD-Nurse* and *Patient Care Assistive Staff*

Figure 5: A screenshot of the browser interface for the Neo4j graph database, showing an example query for a subset of patients in the provider-patient network. Provider nodes are dark blue and patient nodes are blue-green. Directed edges are from providers to patients, indicating that a particular provider accessed a patient's record. Database information including node types, relationship types, and property keys (attributes) are shown on the left side of the figure. The query used to generate the subnetwork is shown at the top of the screen. The black box on the right lists the attributes of the highlighted patient node (circled in gray). This query identifies patients in the data set for whom providers in both the *Dietary 1* and *Rehab PT* positions accessed their record. All protected health information has been deidentified.
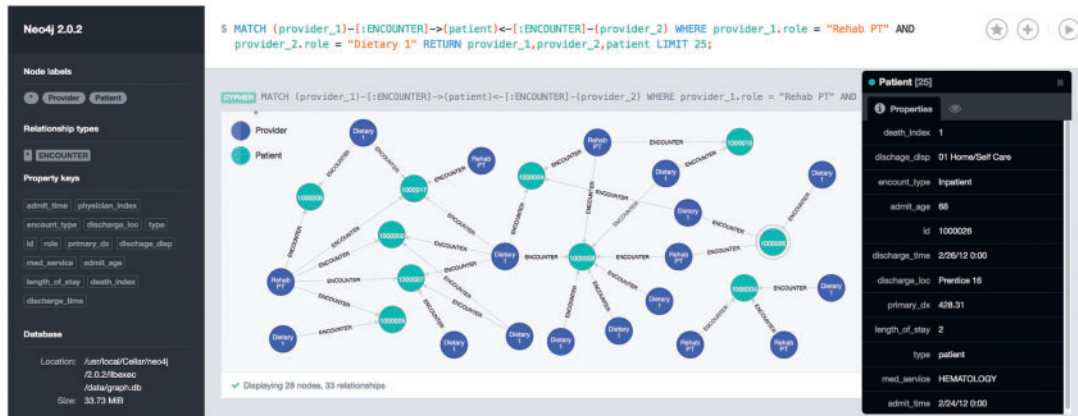
positions had relatively few multidisciplinary collaborations in the ≥ 10 shared patients network (8.5 and 23, respectively). *NMH Physicians* and *RAD-Radiologists* both had a relatively high number of multidisciplinary collaborations per provider in the 10 or more shared patients network and relatively few in the comprehensive network. *Medical Students*, the position that had a relatively high total number of providers (n = 290) and number of collaborations (n = 90 426) in the comprehensive network, had a low number of multidisciplinary collaborations in the ≥ 10 shared patients network (n = 4.5).

Figure 2 illustrates an example of a provider-patient network for a single patient. The pie chart surrounding the network describes the composition of the patient's care team based on each provider's position. The inset shows the provider-patient network with only providers designated as physicians in the Cerner system. This physician-only network contained 24 providers compared to 85 in the comprehensive network. The provider node closest to the patient (position type *Physician*) was involved with placing a total of 58 orders. *Patient Care Staff Nurse* was the position with the most individuals placing orders (n = 4), while the position type *Physician* placed the most orders for this patient. With regards to this comprehensive provider network, administrative and quality assurance staff (position type *Other*) made up 34% (29 of 85) of the patient's care team and had the most patient record accesses. Following *Other, Patient Care Staff Nurse* was the second most represented position, followed in order by *Physician, Resident/Fellow, Patient Care Assistive Staff,* and *Respiratory.* Excluding *Other,* this was roughly consistent with the median number of record accesses by position across the whole network, with the top 5 roles in order being *Patient Care*

Staff Nurse, Respiratory, Resident/Fellow, Patient Care Assistive Staff, and *"Pharmacist"* (table 1).

Using the provider-patient network, we constructed a provider collaboration network by identifying providers who accessed the same patient records (figure 2A). Nodes represent providers, and an edge between 2 nodes indicates that these 2 providers accessed at least 10 common patient records. One thousand five hundred four providers were included, ∼29% (1504 of 5113) of the providers in the initial data set. The total number of edges was 83 998. Nodes were sized according to degree (ie, number of other providers with whom a provider has record accesses in common). Edges were weighted using the shared patient index (see *Materials and Methods* section). The average node degree was approximately 111, indicating a high level of collaboration between providers. Additionally, the average clustering coefficient value (0.882) reflected the strong tendency for pairs of collaborative physicians to also have other collaborative partners in common. However, the network density was quite low (0.074), indicating that there were many more possible collaborative interactions than existed in the real network.

The community detection method identified 7 modules in this network and a network modularity value of 0.256. Figure 3B shows this modular view of provider collaboration with provider nodes grouped by module and a total of 20 edges between them. Edges between modules were weighted by the total number of connections between members of 2 modules. We ordered the modules by the number of providers belonging to each of the following: module 1 (n = 788), module 2 (n = 237), module 3 (n = 219), module 4 (n = 106), module 5 (n = 85), module 6 (n = 43), and module 7 (n = 26) (data not

Table 1: Patient-provider Characteristics by Position

Position	No. of Providers	Patients, No. (%)	Median Providers/Patient, No. (min, max)	No. of Record Accesses	Median Record Accesses, No. (min, max)
Patient Care Staff Nurse	940	543 (99)	12 (0, 123)	9014	4 (1, 264)
NMH Resident/Fellow	654	522 (95)	9 (0, 95)	7000	5 (1, 110)
Med Student	290	272 (50)	0 (0, 9)	622	1 (1, 12)
NMH Physician	262	494 (90)	3 (0, 28)	2397	3 (1, 113)
Patient Care Assistive Staff	238	533 (97)	8 (0, 40)	4662	10 (1, 170)
ED Patient Care Staff Nurse	170	324 (59)	1 (0, 15)	762	2 (1, 62)
Respiratory	107	546 (100)	12 (0, 67)	8761	51 (1, 487)
Student Nurse	101	155 (28)	0 (0, 8)	395	2 (1, 20)
RAD-Technologist	89	335 (61)	1 (0, 11)	501	5 (1, 171)
RX-Pharmacist	84	530 (97)	3 (0, 27)	2567	19 (1, 263)
RAD-Nurse	78	261 (48)	0 (0, 19)	810	7 (1, 48)
RX-Pharmacy Student	77	128 (23)	0 (0, 4)	163	2 (1, 11)
NMH Physician Hospitalist	72	386 (70)	1 (0, 16)	1332	16 (1, 77)
RAD-Resident/Fellow	67	153 (28)	0 (0, 10)	254	3 (1, 12)
Advanced Practice Clinician	61	327 (60)	1 (0, 15)	983	5 (1, 106)
RAD-Radiologist	50	279 (51)	1 (0, 7)	501	3 (1, 57)
NMH Anesthesia	49	109 (20)	0 (0, 14)	313	2 (1, 28)

ED, Emergency department; RAD, Radiology; NMH, Northwestern Memorial Hospital.

shown). Modules 1 and 2 had the most intermodule interactions, while only modules 6 and 7 shared no interactions. Among the 131 provider position types, 74 positions were represented in at least one module. *Patient Care Staff Nurses* were included in each of the 7 modules and they accounted for a high proportion of the providers in modules 3, 5, 6, and 7. Module 1 had the largest number of physicians (n = 234), but interestingly, module 6 had the highest percentage of physicians (33%, 14 of 43). Hospitalists were present in modules 4 and 7 only and represented the largest percentage of any provider in module 4 (26%, 27 of 106). *Respiratory* was represented most prominently in module 2 (15%, 36 of 237). Some modules showed special compositions of positions, while module 1 included at least one representative of most of the provider positions. Module 4 was composed solely of emergency department physicians and nurses, and hospitalists. Moreover, module 6 contained a large proportion of radiology-related providers. Module 5 did not include any physician positions, while module 7 included only 1 physician, an *Oncology Hospitalist*.

The clique detection algorithm was run on the network of healthcare providers sharing ≥ 10 patients. We plotted the number of cliques by the number of providers in each clique in figure 4. A total of 12 355 cliques were identified for 1504 providers. It is important to note that clique membership is not exclusive; a provider may belong to more than one clique. Indeed, this is the case for the majority of the providers in this network (only 45%, 673 of 1504, belonged exclusively to one clique). There were 37 providers who belonged to 10 000 cliques or more. Each of these providers was designated as "non-physician" with their specific position being *Patient Care Staff Nurse, Radiology Technologist, Respiratory, or Pharmacist*. The highest number of clique memberships was among providers in the *Respiratory* position (12 338 cliques). The largest clique included 133 providers, while the average size was 87.9 providers and the median was 96. These results indicate that, in this particular collaboration network, many providers were closely connected, which implies that providers tended to collaborate with large groups of other providers as opposed to collaborating in smaller, tight-knit groups. It can be inferred from this analysis that providers in these positions tended to collaborate with many different groups of providers, illustrating a highly dynamic system of heart failure care. With 45% (673 of 1504) of providers belonging to only one clique, we observed that many tended to collaborate with a constant group of providers, despite what the high average node degree and clustering coefficient may have suggested.

Table 2: Multidisciplinary Patient Record Sharing by Provider Position Type

Position	Any shared patient records			10+ shared patient records		
	No. of Providers	Total No. of Collaborations	Median Collaborations (min, max)	No. of Providers	Total No. of Collaborations	Median collaborations (min, max)
Patient Care Staff Nurse	940	627 026	516 (26, 3580)	265	18 645	33.5 (1, 819)
NMH Resident/Fellow	654	546 923	721.5 (43, 2926)	226	14 544	25 (1, 443)
Med Student	290	90 426	245.5 (24, 1319)	2	9	4.5 (4, 5)
NMH Physician	262	198 237	550 (21, 3374)	62	5867	44 (1, 593)
Patient Care Assistive Staff	238	229 343	832 (26, 3096)	121	9458	23 (1, 474)
ED Patient Care Staff Nurse	170	73 086	261.5 (20, 2230)	18	742	17.5 (1, 152)
Respiratory	107	28 4029	2711 (146, 4960)	92	34 094	283 (3, 1391)
Student Nurse	101	39 627	315 (36, 1284)	13	234	10 (1, 43)
RAD–Technologist	89	74 860	697 (79, 3944)	25	920	4 (1, 756)
RX–Pharmacist	84	137 768	1709.5 (52, 4030)	50	9852	128 (1, 918)
RAD–Nurse	78	85 682	1067.5 (71, 2848)	32	1049	8.5 (1, 239)
RX–Pharmacy Student	77	21 537	216 (37, 1146)	2	9	4.5 (1, 8)
NMH Physician Hospitalist	72	79 627	18 (1, 179)	48	1909	18 (1, 179)
RAD–Resident/Fellow	67	39 835	529 (66, 1605)	4	16	3 (1, 5)
Advanced Practice Clinician	61	63 908	204 (1, 334)	23	3959	204 (1, 334)
NMH Physician Office	50	16 828	238.5 (37, 1290)	1	1	1 (1, 1)
RAD–Radiologist	50	40 405	514.5 (47, 2623)	14	1419	56 (5, 295)
NMH Anesthesia	49	41 016	614 (107, 2328)	11	821	82 (1, 189)

RESEARCH AND APPLICATIONS

We used classification rules generated from a decision tree analysis to query the provider-patient bipartite network using a graph database (supplemental figure S1). The 26 patients who were identified by the rule were associated with a total of 938 providers. There were 2157 interactions between patients and providers, which we termed *encounters* in the database. We derived the following Cypher query from the rule:

```
MATCH (provider_1)-[:ENCOUNTER]->(patient)
 <-[:ENCOUNTER]-(provider_2) WHERE
 provider_1.role = "Rehab PT"
AND provider_2.role = "Dietary 1"
RETURN provider_1,provider_2,patient;
```

This query returned 60 nodes and 107 relationships. Figure 5 shows these results (deidentified and limited to 25 nodes for clarity). The characteristics of 1 patient are shown on the right (accessible by clicking on the node). The Neo4j graph database allowed us to extract information about this specific subset of patients for use in further analyses. Multiple relationship types could be specified as well. For instance, a relationship between providers based on shared patients could be created and displayed within the same query (data not shown).

DISCUSSION

Multidisciplinary collaboration among healthcare professionals occurs daily throughout a hospital stay for patients with heart failure. Our case study of patients with heart failure indicates that healthcare providers have a strong tendency to access healthcare records for common sets of patients, indicating likely collaboration with other providers. This collaboration tends to take place not only in a pairwise manner but also among large groups of providers, resulting in large, multifaceted teams of healthcare providers working implicitly (if not explicitly) together. The complexity of these data sets requires sophisticated and dynamic analysis methods to extract knowledge from the data that will help both cut patient care costs and length of stay while improving the quality of patient care. This study is a first step in strategically guiding the organization of healthcare personnel by network analysis. We now understand the magnitude of actors and the complexity of their interactions.

The current literature on care coordination for heart failure lists a limited but important set of actors necessary for a successful recovery and avoidance of a rehospitalization.[13,43] While each of these actors is important, we found that many more actors remain obscured because their existence is only documented through system audit logs. The use of these logs for security[33,34,38–40] and tracking the use of clinical documentation[37] have been previously described. We contribute a first study evaluating the use of the data as the foundation of an analytic platform for care coordination.

While visualization of the networks is insightful, we feel that our study is limited in several ways. Our EDW contains 131 unique positions, most of which intuitively reveal their function. However, the most frequent positions *Respiratory* and *Patient Care Staff Nurse* could benefit from further differentiation.

Particular providers were members on a very high number of teams (a maximum of 457). Further investigation into this particular provider position or whether this is in fact a shared system login would remove noise from the analysis. In addition, our results do not take into account length of stay. As longer hospital stays appear more multidisciplinary with a greater number of actors, our metrics require further refinement through normalization.

As is often the case in pipeline development, pre- and post-processing of input and output files were necessary to link software packages together. While assembling our data processing pipeline, we found many of the software tools, mostly developed for the digital economy, to be in a nascent state. In addition, while the hardware used in this analysis met or exceeded the requirements to handle our data volume, much of the software we used was not capable of taking full advantage of these resources. As a result, it was necessary to spend a significant amount of time creating a work-around to compensate for these shortcomings. While these tools were highly useful and chosen for their unique functionality and/or their high-quality, publication-level output, their performance will surely benefit from community development and testing over time.

CONCLUSION

EHRs encode valuable interactions, implicitly or explicitly, between patients and providers. Network analysis provides strong evidence of daily multidisciplinary record access of patients with heart failure across teams of 100+ providers. Further investigation may lead to clearer understanding of how record access information can be used to strategically guide care coordination for hospitalizations for heart failure.

FOOTNOTES

SUPPLEMENTARY MATERIAL

Supplementary material is available online at http://jamia.oxfordjournals.org/.

REFERENCES

1. Executive Office of the President Council of Economic Advisers. Reducing Costs and Improving the Quality of

Healthcare. In*: Economic Report of the President (2013)*. http://fraser.stlouisfed.org/publication/?pid=45. Accessed October 1, 2013.

2. Cutler D. Analysis & commentary. How health care reform must bend the cost curve. *Health Aff*. 2010;29(6): 1131–1135.

3. Meltzer DO, Ruhnke GW. Redesigning care for patients at increased hospitalization risk: the comprehensive care physician model. *Health Aff*. 2014;33(5):770–777.

4. Readmissions Reduction Program. Centers for Medicare & Medicaid Services website. http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program.html. Accessed October 1, 2013.

5. Bonow RO, Masoudi FA, Rumsfeld JS, *et al*. ACC/AHA classification of care metrics: performance measures and quality metrics: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. *J Am Coll Cardiol*. 2008;52:2113–2117.

6. Schultz EM, Pineda N, Lonhart J, Davies SM, McDonald KM. A systematic review of the care coordination measurement landscape. *BMC Health Serv Res*. 2013;13:119.

7. Stange KC, Etz RS, Gullett H, *et al*. Metrics for assessing improvements in primary health care. *Annu Rev Public Health*. 2014;35:423–442.

8. Care Coordination Measures Atlas. Agency for Healthcare Research and Quality website. http://www.ahrq.gov/professionals/prevention-chronic-care/improve/coordination/atlas2014/ccm_atlas.pdf. Accessed July 14, 2014.

9. Bonow RO, Ganiats TG, Beam CT, *et al*. ACCF/AHA/AMA-PCPI 2011 performance measures for adults with heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures and the American Medical Association-Physician Consortium for Performance Improvement. *J Am Coll Cardiol*. 2012;59:1812–1832.

10. Fonarow GC, Albert NM, Curtis AB, *et al*. Incremental reduction in risk of death associated with use of guideline-recommended therapies in patients with heart failure: a nested case-control analysis of IMPROVE HF. *J Am Heart Assoc*. 2012;1(1):16–26.

11. Grady KL, Dracup K, Kennedy G, *et al*. Team management of patients with heart failure: a statement for healthcare professionals from The Cardiovascular Nursing Council of the American Heart Association. *Circ*. 2000;102: 2443–2456.

12. Krumholz HM, Brindis RG, Brush JE, *et al*. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. *Circ*. 2006;113: 456–462.

13. Krumholz HM, Currie PM, Riegel B, *et al*. A taxonomy for disease management: a scientific statement from the American Heart Association Disease Management Taxonomy Writing Group. *Circ*. 2006;114:1432–1445.

14. Rudin RS, Bates DW. Let the left hand know what the right is doing: a vision for care coordination and electronic health records. *J Am Med Inform Assoc*. 2014;21(1):13–16.

15. Tai-Seale M, Wilson CJ, Panattoni L, *et al*. Leveraging electronic health records to develop measurements for processes of care. *Health Serv Res*. 2014;49(2):628–644.

16. McAlister FA, Lawson FM, Teo KK, *et al*. A systematic review of randomized trials of disease management programs in heart failure. *Am J Med*. 2001;110:378–384.

17. McAlister FA, Stewart S, Ferrua S, *et al*. Multidisciplinary strategies for the management of heart failure patients at high risk for admission: a systematic review of randomized trials. *J Am Coll Cardiol*. 2004;44:810–819.

18. Sochalski J, Jaarsma T, Krumholz HM, *et al*. What works in chronic care management: the case of heart failure. *Health Aff*. 2009;28:179–189.

19. Peikes D, Chen A, Schore J, Brown R. Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials. *JAMA*. 2009;301(6):603–618.

20. Naylor MD, Aiken LH, Kurtzman ET, Olds DM, Hirschman KB. The care span: the importance of transitional care in achieving health reform. *Health Aff*. 2011;30(4):746–754.

21. Berry LL, Rock BL, Smith Houskamp B, Brueggeman J, Tucker L. Care coordination for patients with complex health profiles in inpatient and outpatient settings. *Mayo Clin Proc*. 2013;88(2):184–194.

22. Hernandez AF, Greiner MA, Fonarow GC, *et al*. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. *JAMA*. 2010;303(17):1716–1722.

23. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. *JAMA*. 2003;289:305–312.

24. Pham HH, O'Malley AS, Bach PB, Saiontz-Martinez C, Schrag D. Primary care physicians' links to other physicians through Medicare patients: the scope of care coordination. *Ann Intern Med*. 2009;150(4):236–242.

25. Bernheim SM, Grady JN, Lin Z, *et al*. National patterns of risk-standardized mortality and readmission for acute myocardial infarction and heart failure. Update on publicly reported outcomes measures based on the 2010 release. *Circ Cardiovasc Qual Outcomes*. 2010;3:459–467.

26. Krumholz HM, Merrill AR, Schone EM, *et al*. Patterns of hospital performance in acute myocardial infarction and heart failure 30-day mortality and readmission. *Circ Cardiovasc Qual Outcomes*. 2009;2:407–413.

27. Goodman DC, Fisher ES, Chang CH, Raymond SR, Bronner KK. *After Hospitalization: A Dartmouth Atlas Report on Post-acute Care for Medicare Beneficiaries*. Lebanon, NH: The Dartmouth Institute for Health Policy & Clinical Practice; 2013.

28. Coleman EA, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's

RESEARCH AND APPLICATIONS

perspective: the care transitions measure. *Med Care*. 2005; 43:246–255.

29. Wee S-L, Loke C-K, Liang C, *et al*. Effectiveness of a national transitional care program in reducing acute care use. *J Am Geriatr Soc*. 2014;62(4):747–753.

30. McLeod J, Stolee P, Walker J, Heckman G. Measuring care transition quality for older patients with musculoskeletal disorders. *Musculoskeletal Care*. 2014;12(1):13–21.

31. Chambers D, Wilson P, Thompson C, Harden M. Social network analysis in healthcare settings: a systematic scoping review. *PLoS One*. 2012;7(8):e41911.

32. Bray-Hall ST. Transitional care: focusing on patient-centered outcomes and simplicity. *Ann Intern Med*. 2012; 157(6):448–449.

33. Gunter CA, Liebovitz DM, Malin B. Experience-based access management: a life-cycle framework for identity and access management systems. *IEEE Secur Privacy.* 2011;9(5): 48–55.

34. Gupta S, Hanson C, Gunter CA, *et al*. Modeling and detecting anomalous topic access. *IEEE Intell Secur Inform.* June 4–7, 2013; Seattle, Washington.

35. Zhang W, Chen Y, Gunter CA, Liebovitz D, Malin B. Evolving role definitions through permission invocation patterns. In: *ACM Symposium on Access Control Models and Technologies*; June 12–14, 2013; Amsterdam, Netherlands.

36. Duffy E, Nyemba S, Gunter CA, Liebovitz D, Malin B. Requirements and design for an extensible toolkit for analyzing EMR audit logs. In: *USENIX Workshop on Health Information Technologies*; August 12, 2013; Washington, D.C.

37. Hripcsak G, Vawdrey DK, Fred MR, Bostwick SB. Use of electronic clinical documentation: time spent and team interactions. *J Am Med Inform Assoc*. 2011;18(2):112–117.

38. Martin P, Rubin AD. Enforcing minimum necessary access in healthcare through integrated audit and access control. In: *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics*. September 22–25, 2013; Washington DC. 946–955.

39. Fabbri D, Lefevre K. Explaining accesses to electronic medical records using diagnosis information. *JAMIA*. 2013; 20(1):52–60.

40. Chen Y, Lorenzi N, Nyemba S, Schildcrout JS, Malin B. We work with them? Healthcare workers interpretation of organizational relations mined from electronic health records. *Int J Med Inform*. 2014;83(7):495–506.

41. Coleman EA, Boult C. Improving the quality of transitional care for persons with complex care needs. *J Am Geriatr Soc*. 2003;51:556–557.

42. Salmon RB, Sanderson MI, Walters BA, *et al*. A collaborative accountable care model in three practices showed promising early results on costs and quality of care. *Health Aff*. 2012;31(11):2379–2387.

43. Lambrinou E, Kalogirou F, Lamnisos D, Sourtzi P. Effectiveness of heart failure management programmes with nurse-led discharge planning in reducing re-admissions: a systematic review and meta-analysis. *Int J Nurs Stud*. 2012;49(5):610–624.

44. Munoz DA, Alonso W, Nembhard HB. A social network analysis-based approach to evaluate workflow and quality in a pediatric intensive care unit. In: *Proceedings of the 2014 Industrial and Systems Engineering Research Conference.* May 31-June 3, 2014, Montreal, Canada.

45. Anderson C, Talsma AN. Characterizing the structure of operating room staffing using social network analysis. *Nurs Res*. 2011;60(6):378–385.

46. Gray JE, Davis DA, Pursley DM, *et al*. Network analysis of team structure in the neonatal intensive care unit. *Pediatrics*. 2010;125(6):e1460–e1467.

47. Bosch M, Faber MJ, Cruijsberg J, *et al*. Review article: effectiveness of patient care teams and the role of clinical expertise and coordination: a literature review. *Med Care Res Rev*. 2009;66:5S–35S.

48. Luke DA, Harris JK. Network analysis in public health: history, methods, and applications. *Annu Rev Public Health*. 2007;28:69–93.

49. Uzzi B. A social network's changing statistical properties and the quality of human innovation. *J Phys A: Mathe Theor*. 2008;41(22):224023.

50. Newman MEJ. The structure of scientific collaboration networks. *Proc Natl Acad Sci USA*. 2001;98(2): 404–409.

51. Pan RK, Kaski K, Fortunato S. World citation and collaboration networks: uncovering the role of geography in science. *Sci Rep*. 2012;2:902.

52. Newman MEJ. *Networks: An Introduction*. New York, NY: Oxford University Press; 2010.

53. Microsoft. Microsoft SQL Server Management Studio 2012 http://www.microsoft.com/en-us/download/details.aspx?id=29062. Accessed January 15, 2014.

54. The Perl Programming Language. http://www.perl.org Accessed April 1, 2014.

55. The Python Language. https://www.python.org. Accessed April 1, 2014.

56. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In*: International AAAI Conference on Weblogs and Social Media*. May 17–20, 2009; San Jose, California.

57. R: RCoreTeam. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2012.

58. Neo Technology. The Neo4j Manual v2.0.2. http://www.neotechnology.com. Accessed April 1, 2014.

59. Jaccard, P. Nouvelles recherches sur la distribution florale. *Bull Soc Vaudoise Sci Nat*. 1908;44:223–270.

60. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. *Software: Pract Exper*. 1991;21(11): 1129–1164.

61. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks*. J Stat Mech Theory Exp*. 2008;10:1000.

62. Wasserman S, Faust K. *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University; 1994:258.

63. Carey V, Long L, Gentleman R. RBGL: An interface to the BOOST graph library. R package version 1.38.0. http://www.bioconductor.org. Accessed on April 1st, 2014.

64. Robinson I, Webber J, Eifrem E. *Graph Databases*. Sebastopol, CA: O'Reilly Media; 2013.

65. Therneau TM, Atkinson EJ, Ripley B. An introduction to recursive partitioning using the RPART routines. Version 4.1.8. The Comprehensive R Archive Network website.

http://cran.r-project.org/web/packages/rpart. Accessed May 5, 2014.

66. Hall M, Frank E, Holmes G, *et al*. The WEKA data mining software: an update. *SIGKDD Explorations*. 2009;11(1): 10–18.

67. Breiman L, Friedman J, Olshen R, Stone C. *Classification and Regression Trees*. Pacific Grove: Wadsworth; 1984.

68. Steinberg D, Colla P. *CART: Tree-Structured Non-Parametric Data Analysis*. San Diego, CA: Salford Systems; 1995.

AUTHOR AFFILIATIONS

[1]Division of Health and Biomedical Informatics, Department of Preventive Medicine, Northwestern University, Chicago, IL, USA

[2]Center For Healthcare Studies, Institute for Public Health and Medicine, Northwestern University, Chicago, IL, USA

[3]Northwestern University Clinical and Translational Sciences Institute, Chicago, IL, USA

[4]Division of Biostatistics, Department of Preventive Medicine, Northwestern University, Chicago, IL, USA

RESEARCH AND APPLICATIONS