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

Nicholas D Soulakis¹, Matthew B Carson¹,², Young Ji Lee¹, Daniel H Schneider³, Connor T Skeehan¹, Denise M Scholtens⁴

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.

Key words: Care collaboration, heart failure, electronic health records, network analysis

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 evaluation³² and furthermore lack comprehensive ascertainment of realized care coordination at...
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. 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.

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. 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.

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 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 and Python to create input files for Gephi and Neo4j 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. The second network is undirected and depicts shared patient record access between providers. Visualization for both networks was performed using Gephi. 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
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, 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. This force-directed algorithm considers edge weights in determining the strength of a relationship between 2 nodes. A stronger...
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. 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 al61 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
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 kClique algorithm, which is part of the Boost Graph Library interface package (RBGL) for the R statistical computing environment.

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. 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, WEKA: J48 decision tree, and Salford Predictive Suite: CART).

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...
providers/patients \((\text{min, max})\) 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 \((\text{min, max})\) indicates the median, minimum, and maximum of the previous value. The following top 5 most frequently observed provider positions made up 46% \((2384 \text{ of } 5113)\) of the total positions: Patient Care Staff Nurse \((n = 940; 18\%, 940 \text{ of } 5113)\), NMH Resident/Fellow \((n = 654; 13\% 654 \text{ of } 5113)\), Medical Student \((n = 290; 6\%, 290 \text{ of } 5113)\), NMH Physician \((n = 262; 5\%, 62 \text{ of } 5113)\), and Patient Care Assistive Staff \((n = 238; 5\%, 238 \text{ of } 5513)\). 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% \((32,004 \text{ of } 61,026)\) of these record accesses: Patient Care Staff Nurse \((n = 9014; 15\%, 9014 \text{ of } 61026)\), Respiratory \((n = 8761; 14\%, 8761 \text{ of } 61026)\), NMH Resident/Fellow \((7000; 11\%, 7000 \text{ of } 61026)\), Patient Care Assistive Staff \((n = 4662; 8\%, 4662 \text{ of } 61026)\), and Rx-Pharmacist \((n = 2567; 4\%, 2567 \text{ of } 61026)\).

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 Anesthesia \((n = 82)\). For the collaboration network with \(\geq 10\) shared patients, the top 5 positions in this respect were Respiratory \((n = 283)\), Advanced Practice Clinician \((n = 204)\), RX-Pharmacist \((n = 128)\), Patient Care Assistive Staff \((n = 62)\), 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 made up a significant portion of these collaborations. The Respiratory position had the largest number of collaborations, but each provider interacted with a median of 12 patients. The number of collaborations per provider is a measure of the level of interaction between providers in the network, and the median number of collaborations per provider gives an indication of the average level of interaction for each position type.
positions had relatively few multidisciplinary collaborations in the \(\geq 10\) 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 \(\geq 10\) 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 “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, \(\sim 29\%\) (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...
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%, 36 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 1: Patient-provider Characteristics by Position

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

ED, Emergency department; RAD, Radiology; NMH, Northwestern Memorial Hospital.
<table>
<thead>
<tr>
<th>Position</th>
<th>No. of Providers</th>
<th>Total No. of Collaborations</th>
<th>Median Collaborations (min, max)</th>
<th>No. of Providers</th>
<th>Total No. of Collaborations</th>
<th>Median collaborations (min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Care Staff Nurse</td>
<td>940</td>
<td>627 026</td>
<td>516 (26, 3580)</td>
<td>265</td>
<td>18 645</td>
<td>33.5 (1, 819)</td>
</tr>
<tr>
<td>NMH Resident/Fellow</td>
<td>654</td>
<td>546 923</td>
<td>721.5 (43, 2926)</td>
<td>226</td>
<td>14 544</td>
<td>25 (1, 443)</td>
</tr>
<tr>
<td>Med Student</td>
<td>290</td>
<td>90 426</td>
<td>245.5 (24, 1319)</td>
<td>2</td>
<td>9</td>
<td>4.5 (4, 5)</td>
</tr>
<tr>
<td>NMH Physician</td>
<td>262</td>
<td>198 237</td>
<td>550 (21, 3374)</td>
<td>62</td>
<td>5867</td>
<td>44 (1, 593)</td>
</tr>
<tr>
<td>Patient Care Assistive Staff</td>
<td>238</td>
<td>229 343</td>
<td>832 (26, 3096)</td>
<td>121</td>
<td>9458</td>
<td>23 (1, 474)</td>
</tr>
<tr>
<td>ED Patient Care Staff Nurse</td>
<td>170</td>
<td>73 086</td>
<td>261.5 (20, 2230)</td>
<td>18</td>
<td>742</td>
<td>17.5 (1, 152)</td>
</tr>
<tr>
<td>Respiratory</td>
<td>107</td>
<td>28 4029</td>
<td>2711 (146, 4960)</td>
<td>92</td>
<td>34 094</td>
<td>283 (3, 1391)</td>
</tr>
<tr>
<td>Student Nurse</td>
<td>101</td>
<td>39 627</td>
<td>315 (36, 1284)</td>
<td>13</td>
<td>234</td>
<td>10 (1, 43)</td>
</tr>
<tr>
<td>RAD-Technologist</td>
<td>89</td>
<td>74 860</td>
<td>697 (79, 3944)</td>
<td>25</td>
<td>920</td>
<td>4 (1, 756)</td>
</tr>
<tr>
<td>RX-Pharmacist</td>
<td>84</td>
<td>137 768</td>
<td>1709.5 (52, 4030)</td>
<td>50</td>
<td>9852</td>
<td>128 (1, 918)</td>
</tr>
<tr>
<td>RAD-Nurse</td>
<td>78</td>
<td>85 682</td>
<td>1067.5 (71, 2848)</td>
<td>32</td>
<td>1049</td>
<td>8.5 (1, 239)</td>
</tr>
<tr>
<td>RX-Pharmacy Student</td>
<td>77</td>
<td>21 537</td>
<td>216 (37, 1146)</td>
<td>2</td>
<td>9</td>
<td>4.5 (1, 8)</td>
</tr>
<tr>
<td>NMH Physician Hospitalist</td>
<td>72</td>
<td>79 627</td>
<td>18 (1, 179)</td>
<td>48</td>
<td>1909</td>
<td>18 (1, 179)</td>
</tr>
<tr>
<td>RAD-Resident/Fellow</td>
<td>67</td>
<td>39 835</td>
<td>529 (66, 1605)</td>
<td>4</td>
<td>16</td>
<td>3 (1, 5)</td>
</tr>
<tr>
<td>Advanced Practice Clinician</td>
<td>61</td>
<td>63 908</td>
<td>204 (1, 334)</td>
<td>23</td>
<td>3959</td>
<td>204 (1, 334)</td>
</tr>
<tr>
<td>NMH Physician Office</td>
<td>50</td>
<td>16 828</td>
<td>238.5 (37, 1290)</td>
<td>1</td>
<td>1</td>
<td>1 (1, 1)</td>
</tr>
<tr>
<td>RAD-Radiologist</td>
<td>50</td>
<td>40 405</td>
<td>514.5 (47, 2623)</td>
<td>14</td>
<td>1419</td>
<td>56 (5, 295)</td>
</tr>
<tr>
<td>NMH Anesthesia</td>
<td>49</td>
<td>41 016</td>
<td>614 (107, 2328)</td>
<td>11</td>
<td>821</td>
<td>82 (1, 189)</td>
</tr>
</tbody>
</table>
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 = "Respiratory"
  AND provider_2.role = "Dietary"
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. 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 and tracking the use of clinical documentation 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

Contributors NDS: drafting of the manuscript. MBC: analysis of data and drafting of the manuscript. YJL: analysis of data and drafting of the manuscript. CTS: analysis of data and drafting of the manuscript. DMS: reviewing of data analysis approaches and editing of the manuscript.

Funding The project described was supported by the National Library of Medicine, Grant K01LM011973, and the National Center for Research Resources, Grant 5UL1RR025741, and is now at the National Center for Advancing Translational Sciences, Grant 8UL1TR000150.

Competing interests None.

Provenance and peer review

Ethics Approval Ethics approval was provided by the Institutional Review Board of Northwestern University.

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


AUTHOR AFFILIATIONS

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

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

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

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