Journal of the American Medical Informatics Association, 0(0), 2018, 1–10
doi: 10.1093/jamia/ocy009
Research and Applications

Research and Applications

Interaction patterns of trauma providers are associated with length of stay

You Chen,[1] Mayur B Patel,[2,3,4] Candace D McNaughton,[5] and Bradley A Malin[1,6,7]

[1]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA, [2]Department of Surgery, Division of Trauma, Surgical Critical Care, and Emergency General Surgery, Vanderbilt University Medical Center, Nashville, TN, USA, [3]Department of Neurosurgery, Vanderbilt University Medical Center, Nashville, TN, USA, [4]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA, [5]Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, TN, USA, [6]Department of Biostatistics, School of Medicine, Vanderbilt University, Nashville, TN, USA and [7]Department of Electrical Engineering and Computer Science, School of Engineering, Vanderbilt University, Nashville, TN, USA

Corresponding Author: You Chen, Department of Biomedical Informatics, Vanderbilt University Medical Center, 2525 West End Ave, Suite 1475, Nashville, TN 37203 USA. E-mail: you.chen@vanderbilt.edu

Received 6 October 2017; Revised 22 December 2017; Editorial Decision 19 January 2018; Accepted 26 January 2018

ABSTRACT

Background: Trauma-related hospitalizations drive a high percentage of health care expenditure and inpatient resource consumption, which is directly related to length of stay (LOS). Robust and reliable interactions among health care employees can reduce LOS. However, there is little known about whether certain patterns of interactions exist and how they relate to LOS and its variability. The objective of this study is to learn interaction patterns and quantify the relationship to LOS within a mature trauma system and long-standing electronic medical record (EMR).

Methods: We adapted a spectral co-clustering methodology to infer the interaction patterns of health care employees based on the EMR of 5588 hospitalized adult trauma survivors. The relationship between interaction patterns and LOS was assessed via a negative binomial regression model. We further assessed the influence of potential confounders by age, number of health care encounters to date, number of access action types care providers committed to patient EMRs, month of admission, phenome-wide association study codes, procedure codes, and insurance status.

Results: Three types of interaction patterns were discovered. The first pattern exhibited the most collaboration between employees and was associated with the shortest LOS. Compared to this pattern, LOS for the second and third patterns was 0.61 days ($P=0.014$) and 0.43 days ($P=0.037$) longer, respectively. Although the 3 interaction patterns dealt with different numbers of patients in each admission month, our results suggest that care was provided for similar patients.

Discussion: The results of this study indicate there is an association between LOS and the extent to which health care employees interact in the care of an injured patient. The findings further suggest that there is merit in ascertaining the content of these interactions and the factors that induce these differences in interaction patterns within a trauma system.

Key words: Interaction network, electronic medical records, length of stay, network analysis, statistical modeling, trauma

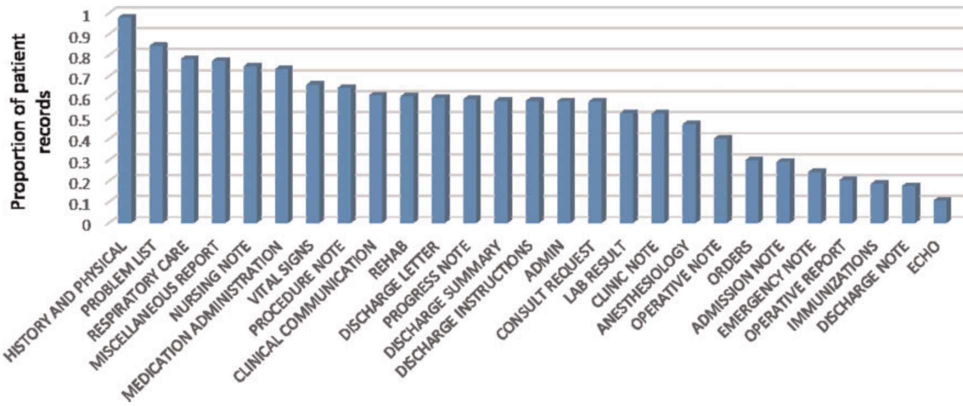

Figure 1. The access action types that were invoked for at least 10% of patient records.

INTRODUCTION

Health care spending continues to escalate in the United States. Expenditures reached \$3.2 trillion, or \$9990 per person, in 2015 and \$3.35 trillion, or \$10 348 per person, in 2016.[1] The rising cost of hospitalizations, which accounted for ~32% of expenditures in 2016, is one of the major driving factors behind higher health care payments.[1] In particular, trauma-related hospitalization has the highest expense in the United States,[2,3] which brings about heavy financial burdens for health care systems, patients, and health insurance companies.[4]

The expenditure for a hospitalization is directly related to the quantity of resources consumed. Notably, a patient's length of stay (LOS) is a key indicator of inpatient resource consumption,[5–7] which is typically measured as the number of days the patient occupies a bed in the hospital. Though it should be recognized that LOS is not the sole indicator of resource consumption, it can serve as a good proxy to characterize the degree to which inpatient resources are consumed.[8,9]

At the same time, it has been recognized that establishing a team in clinical care settings (eg, trauma) can significantly reduce in-hospital mortality and LOS.[10–13] Trauma care often involves interactions among a multidisciplinary group of health care employees (eg, anesthesiologists, surgeons, emergency room physicians, respiratory therapists, nurse practitioners, radiographers, neurosurgeons, and various types of nurses) who are distributed across time and space.[14] Given the high heterogeneity of health care employees who interact with trauma patients, the wide variability in outcomes, and the substantial contribution to financial burden, we focus on this section of the hospital population. We anticipate that quantifying the differences in LOS between trauma patients affiliated with different patterns of health care employee interactions can provide evidence for health care organizations to (1) further investigate the factors leading to such patterns and, ultimately, (2) refine or influence interaction structures in a manner that reduces clinical resource consumption.

Toward realizing this goal, various auditing (eg, video review, observer review, and medical note review)[14,15] and simulation (eg, using simulators who educate team members on communication, cooperation, and leadership)[16] programs have been proposed to assess and refine interaction processes to reduce inpatient resource consumption. These approaches have traditionally adopted an expert-driven management strategy to observe interaction routines and assess their relationship with outcomes. Such strategies, however, often require a nontrivial amount of manual effort to gather the necessary observations and perform the assessment.[17–20]

Given the limitations of existing approaches, a desirable alternative is to develop data-driven strategies, which automatically infer the interaction patterns of employees and characterizes their relationships with respect to patient outcomes. We believe that electronic medical record (EMR) systems can supply the data necessary to support such a strategy. This is because EMR systems capture information sharing, coordination, and documentation longitudinally. As a result, they can provide insight into many operational activities from a diverse collection of health care employees.[21–27] This type of data has shown promise for inferring health care organizational patterns[28,29] and analyzing patient outcomes.[30] Thus, in this study, we introduce such an approach to automatically investigate the relationship between interaction patterns and LOS through data inherent in EMR systems.

METHODS

Study materials

This study focuses on patients who were assigned to the trauma service of Vanderbilt University Medical Center (VUMC) and completed an inpatient stay between December 2013 and December 2015. Annually, VUMC provides state- and national-verified Level 1 trauma care for a geographic region spanning 70 000 square miles and 4 states.[31] Our study leverages the VUMC EMR system to learn trauma-centered interaction patterns of health care employees and quantify their association with hospital LOS. Information associated with patients' hospitalization, as well as health care employees' utilization of such information, is documented in a homegrown EMR system that has been central to clinical activities since the 1990s.[32]

During this period, 5547 employees, affiliated with 179 operational areas in the medical center (eg, mental health center, neuro intensive care unit, and neurosurgery clinic), committed EMR access actions during 5588 patient encounters with the health care system. This entailed 158 467 unique actions and 67 distinct access action types (eg, operative note, physical and history, clinical communication, lab results, vital signs, and medication administration), which were relied upon to infer interaction patterns among health care employees with patients. The access action types that were invoked with respect to at least 10% of the patients are depicted in Figure 1. There were 27 access action types (40%) that met this criterion.

To assess the relationships among interaction patterns and patient outcomes, we extracted hospital LOS for each patient encounter, where LOS is measured as the duration between admission and discharge. We excluded 219 patient encounters where patients died while in trauma care, to focus on interaction patterns of survivors indicative of the completion of hospital care. We neglect this subpopulation because they did not complete a stay at the hospital, and thus their LOS is inaccurate. For instance, some patients died while in transit or shortly after admission to the hospital. Given that LOS may also be related to additional confounders (eg, patient age, degree of illness, procedure burden, history of service utilization, access action type, admission season, and type of health insurance), we extracted such factors for each patient encounter for further investigation. Summary information for these confounders are provided in Table 1.

The EMRs for the patient encounters in this study contained 3612 distinct International Classification of Diseases, Ninth Version (ICD-9) billing codes, 1627 distinct Current Procedural Terminology (CPT) codes, and 8 insurance programs (eg, Medicaid and Medicare Part A).

We acknowledge that ICD-9 codes are not sufficient to accurately represent patient illness. Thus, to mitigate bias due to variable insurance billing, we rely on the phenome-wide association study (PheWAS) vocabulary, which was introduced to group ICD-9 codes together and reduce variability in the definitions of clinical concepts in secondary data use scenarios.[33,34] Upon translating each ICD-9 code, the data consisted of 1010 PheWAS codes.

Study design

We constructed cohorts by grouping patient encounters according to interaction patterns. This was accomplished by (1) inferring the interaction patterns, and then (2) applying the inferred patterns to compose the cohorts. In doing so, patient encounters in each group shared similar interaction patterns, as described below.

Our study design consists of 3 components: (1) learn patient encounter groups according to interaction patterns; (2) quantify interaction patterns according to standard social networking characteristics; and (3) assess the relationships interaction patterns and hospital LOS.

Grouping patients by interaction patterns

We use a binary matrix A to represent health care employees' committing access actions to EMRs. (The matrix is binary because the number of accesses to a particular patient can be artificially inflated due to system design. For instance, if a user accesses a patient's medical record, such as a laboratory report, the system may record the access action multiple times.) Specifically, the value of a cell $A(i,j)$ is 1 if a health care professional i committed an action to the EMR during patient encounter j and 0 otherwise. We leverage information in A to derive interaction patterns and patient encounter groups. It has been shown that interaction patterns inferred at a level of operational health care areas are more stable and interpretable than patterns learned at the level of health care employees.[28,35] Thus, we transform A into a matrix A', where each cell stores the number of actions that all health care employees from a specific operational area committed to the EMR during a patient encounter.

Since the EMR of a patient encounter is typically worked on by a small subset of health care employees, A' is a sparse matrix. Thus, we apply a spectral co-clustering model to A' to uncover groups of patient encounters according to their interaction patterns.[36–38] This methodology employs matrix decomposition techniques and formalizes co-clustering as a bipartite graph partitioning problem.[14,39] The details of the patient encounter grouping process are in Supplementary data S1. We rely on this method because it has been shown to be robust in high-dimensional sparse matrices,[39] which is indicative of our setting (ie, only a portion of the operational areas interact with one another during a patient encounter). All algorithms were implemented in Matlab 2017a.

Quantifying the characteristics of interaction patterns

To ascertain whether interaction patterns are associated with hospital LOS, we represent each interaction pattern as a network of operational areas and then quantify the networks via social network characteristics. For each group of patient encounters, we infer a network of operational areas to show how affiliated health care employees interacted with one another in these encounters. Each node in a network corresponds to an operational area, and each edge weight between 2 operational areas is the cosine similarity of interactions with the EMRs during patient encounters among health care employees from these 2 operational areas.[40] The details for the edge weighting process are in Supplementary data S1.

We leverage standard social networking characteristics to quantify each interaction pattern. Specifically, these characteristics correspond to average node degree, average weighted node degree, graph density, clustering coefficient, and average path length.[41,42] The definitions of these characteristics are as follows:

- Average node degree: Calculated by summing the degree of each node (ie, the number of edges connected to it) and dividing by the total number of nodes.
- Average weighted node degree: Calculated by summing the weighted degree of a node (ie, the sum of weights of edges connected to it) and dividing by the total number of nodes.
- Graph density: The ratio of the number of edges observed to the number of possible edges.
- Clustering coefficient: The average clustering coefficient for all nodes. The cluster coefficient of a node is the ratio of existing edges connecting a node's neighbors to each other to the maximum possible number of such edges. A large clustering coefficient is an indication of high collaboration among employees in a network.
- Average path length: Calculated by summing shortest path lengths between all pairs of nodes and dividing by the total number of pairs. This indicates the number of steps, on average, it takes to move from one node in the network to another.

Subnetworks may exist within each interaction network, so we further infer communities of health care employees in each interaction network. This is accomplished through an algorithm that optimizes the modularity of a network.[43] We guide the algorithm using a heuristic based on optimization of the modularity measure,[43] which is efficient (in running time) and effective (in quality of communities) for weighted and undirected graphs. Modularity is defined as:

$$Q = \frac{1}{2m}\sum_{vw}\sum_{r}[A_{vw} - \frac{k_v k_w}{2m}]S_{vr}S_{wr}, \quad (1)$$

where m is the number of edges in the network, $k_v,\ k_w$ is the degree of vertices v and w, respectively, $A_{vw} = 1$ means there is an edge between the 2 vertices, and S_{vr} is defined as 1 if vertex v belongs to group r and zero otherwise. A community with high modularity has dense connectivity between operational areas within the community, but sparse connectivity between the other communities. We use approaches implemented in the Gephi software suite to quantify standard network characteristics and infer communities.[48]

Assessing the relationship between interaction patterns and LOS

We apply a generalized linear regression model with negative binomial distributions of LOS[44] to test the associations between interaction patterns and LOS. The negative binomial distribution

Table 1. A comparison of the 3 discovered interaction patterns with respect to various control factors considered in this study.

Items	Patient group		
	P_1	P_2	P_3
	($n = 428$)	($n = 1353$)	($n = 3807$)
Standard network characteristics			
Number of operational areas	102	138	125
Degree average	27.14	23.38	22.47
Weighted degree average	7.02	5.78	5.31
Graph density	0.27	0.17	0.18
Cluster coefficient average	0.77	0.70	0.73
Path length average	1.43	2.16	1.52
Characteristics of outcome (LOS) and confounders			
Median LOS (days) (Q1, Q3, IQR)	4.62 (3.19, 7.79, 4.6)	7.12 (2.75, 10.21, 7.46)	6.72 (2.89, 8.99, 6.10)
Median no. of encounters to date (Q1, Q3, IQR)	5 (2, 12, 10)	10 (4, 23, 19)	6 (3, 14, 11)
Median no. of PheWAS codes (Q1, Q3, IQR)	10 (5, 21, 16)	10 (5, 20, 15)	3 (1, 14, 13)
Median no. of CPT codes (Q1, Q3, IQR)	19 (9, 36, 27)	17 (9, 33, 24)	3 (1, 18, 17)
Median no. of access action types (Q1, Q3, IQR)	23 (19, 26, 7)	34 (21, 26, 5)	23 (20, 26, 6)
Median age (Q1, Q3, IQR)	44.8 (30.2, 61.1, 30.9)	48.9 (29.0, 64.5, 35.5)	46.1 (28.5, 62.2, 33.7)
Distribution of admission month (%)			
January	0.47	7.76	6.04
February	2.34	5.32	6.28
March	1.63	3.99	7.30
April	4.67	5.91	7.85
May	23.1	5.17	7.30
June	27.3	5.69	7.38
July	13.6	5.32	7.14
August	9.8	4.51	10
September	3.7	8.06	10.16
October	4.9	13.01	11.74
November	4.4	19.59	9.19
December	3.9	15.67	9.59
Distribution of insurance programs (%)			
Commercial	29.2	29.9	28.3
Blue Cross	16.8	16.5	16
Medicare Part A	14.5	18.8	14.9
Medicaid	14.5	11.1	14
HMO	2.1	1.1	2.1
Champus	2.1	0.96	1.4
Medicare Part B	0.47	0.74	0.26
Unknown	20.3	21	23

CPT = Current Procedural Terminology; HMO = Health Maintenance Organization; PheWAS = phenome-wide association study

has been shown to achieve the best performance for normalizing hospital LOS, whose distribution does not follow a normal distribution.[45] We add variables for interaction patterns, age, number of encounters to date, number of access action types, month of admission, insurance programs, and health conditions into the regression to estimate the coefficients of the model, and then leverage the fitted model to derive adjusted LOS for each specific interaction pattern. We applied analysis of variance[46] with a 95% confidence interval to test the significance of differences in LOS for pairs of interaction patterns. We relied on standard packages in Matlab to compute the generalized linear regression model and analysis of variance.

We further investigated whether the differences in LOS were correlated with the potential confounding factors that we incorporated in the regression models. This was accomplished by testing for differences in the distributions of PheWAS codes, procedure codes, insurance programs, age, number of health care encounters to date, admission month, and access action types between each pair of patient encounter groups. Specifically, for each pair of patient encounter groups, we compare the similarity in the distributions of the aforementioned factors through a Pearson correlation coefficient (PCC).[47] This similarity score is in the range (−1,1), where 1 indicates a positive direct correlation, 0 indicates no correlation, and −1

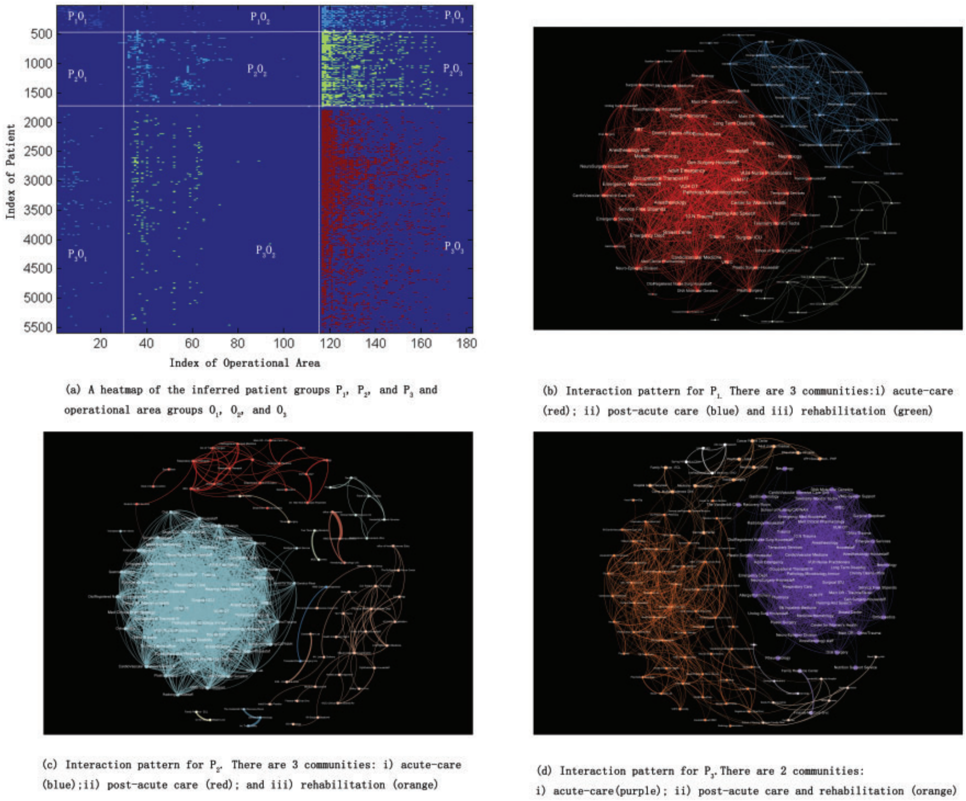

Figure 2. (**A**) A heatmap of the inferred patient groups P1 (428 patients), P2 (1353), and P3 (3807) and operational area groups O1 (27 areas), O2 (86 areas), and O3 (66 areas). (**B–D**) Three interaction patterns for P1 through P3. Each pattern is composed of operational health care areas coming from all 3 groups, but with different network structures. Each interaction pattern is associated with 3 types of care (1) acute-care, (2) post-acute care, and (3) rehabilitation.

indicates a negative direct correlation. If each pair of patient encounter groups exhibits a high PCC with significance at the 95% confidence level for each factor, then we consider them to be sufficiently similar. If a pair of groups exhibits similar distributions in terms of these factors, then their corresponding interaction patterns likely handle a similar patient population. This would lend credibility to the claim that the different interaction patterns manage similar patients, but with different hospital LOS. Further details about this assessment are in Supplementary data S2.

RESULTS

Patient encounter groups and interaction patterns

The co-clustering approach discovered 3 patient encounter groups, which we refer to as P_1, P_2, and P_3. These groups were composed of 428, 1353, and 3807 inpatient encounters, respectively. Additionally, the approach discovered 3 operational groups, which we refer to as O_1, O_2, and O_3. The relationship between the patient encounter and operational groups is depicted as a heatmap in Figure 2 (A). In this figure, each point indicates the number of actions that health care employees from a certain operational area committed to the EMR during patient encounters. For each patient encounter group, it can be seen that all 3 operational groups are involved, but with different interaction patterns, as shown in Figure 2 (B–C) for P_1, P_2, and P_3.

Each interaction pattern is composed of 3 major communities, which are informally characterized as: (1) acute care team, including representative operational areas such as “emergency,” “anesthesiology,” and “cardiovascular intensive care unit”; (2) post-acute care team, including operational areas such as “nutrition clinic,” “respiratory therapist,” and “social work”; and (3) rehabilitation team, including operational areas such as “rehabilitation service” and “physical medicine and rehab department.” It can be seen that acute care achieved a high density of collaboration in all 3 patterns, but particularly so in P_1. However, there are notably large differences in the network structures of the other 2 communities across the 3 interaction patterns. For instance, in the second pattern, the network structures of post-acute care (red) and rehabilitation (orange) are very sparse, and in the third pattern, the post-acute care and rehabilitation communities overlapped.

Quantified interaction patterns

The quantified network characteristics for the interaction patterns are reported in Table 1. It can be seen that the interaction pattern for patient group P_1 was affiliated with the smallest number of operational areas (102). However, at the same time, this interaction pattern realized the highest amount of collaboration among employees from these operational areas (an average degree of 27.14, an average weighted degree of 7.02, a graph density of 0.27, a cluster coefficient of 0.77, and an average path length of 1.43).

The results in Table 1 indicate several notable findings with respect to interactions among health care employees. First, the interaction pattern for P_1 has the highest average degree (in comparison to interaction patterns for P_2 and P_3), which indicates that health care employees in this pattern establish more connections with other

Table 2. Differences in operational areas between the networks of the 3 patient groups.

Patient groups	Operational areas in the first group but not in the second group	Operational areas in the second group but not in the first group
P_1 vs P_2	Eskind Biomed Library; Neurology; School of Nursing Academic Faculty; Technical Systems Infrastructure; Vanderbilt Medical Group Executive Leadership; Vanderbilt Orthopaedic Institute – Rehab Services	Cardiovascular Intensive Care Unit; Briley Sterile Processing; Cancer Patient Center; Cardiac Catheterization Lab; Center for Human Genetics Research; Clinical Pharmacology; Clinical Staffing Resource Center; Center for Surgical Weight Loss; Evolve to Excel – Hospital and Clinic; Health Information Management; International Travel Clinic; Medical Center East Recovery; Point of Care Testing; Psychiatry Reference Lab; Respiratory Care; Registered Nurse Center; Thoracic Surgery; Clinical Toxicology and Therapeutic Drug Monitoring; Transplant Center Admin; Vanderbilt Medical Group Claims Management; Vanderbilt Medical Group Expected Reimbursement; Vanderbilt Orthopaedic Institute Radiology; Vanderbilt Network Services; Office of Health Sciences Education; Office of Research
P_1 vs P_3	Department of Oral Surgery; Department of Surgery Admin; Division of Trauma Surgeon; Eskind Biomed Library; Hematology Oncology Unit; Registered Nurse Medicine; Main Operation Room – Post-Acute Care Unit; Nuclear Lab; Nutrition Clinic; Pathology; Respiratory Therapist; School of Nursing Academic Faculty; Social Work; Special Education; Student Health Services; Surgery; Technical Systems Infrastructure; Transplant/General Surgery Unit; Vanderbilt University Non-Employee Payments; Walk-in Clinics–Hwy 96	Cardiovascular Intensive Care Unit; Briley Sterile Processing; Cancer Patient Center; Cardiac Catheterization Lab; Center for Human Nutrition; Clinical Staffing Resource Center; Vanderbilt Multiple Sclerosis Center; Department of Pharmacology; Evolve to Excel – Hospital and Clinic; Family Resource Center; Health Information Management; Heart Institute; Heart Station-EKG; Hospital Guest Services; Informatics Center; Registered Nurse Medicine; Kennedy Center; Medical Center East Recovery; Medicine – Rheumatology; Mental Health Center; Myelosuppression; Neuro – Epilepsy Lab; Palliative Care/Inpatient Med; School of Nursing Control Center; Psychiatry and Neurology; Radiology Administration; Reference Lab; Rehab; Student Workers; Respiratory Care; Rheumatology Infusion; School of Nursing Academic Faculty Other; School of Nursing Research Admin; Sleep Lab; Specimen Receiving; Spring Hill Walk-in Clinic; Thoracic Surgery; Vanderbilt Orthopaedic Institute Radiology; Vanderbilt Psychiatric Hospital Social Work; Vanderbilt University Police Department; Vanderbilt Network Services; Office of Health Sciences Education
P_2 vs P_3	Center for Human Genetics Research; Clinical Pharmacology; Center for Surgical Weight Loss; Department of Oral Surgery; Department of Surgery Admin; Division of Trauma Surgeon; Hematology Oncology Unit; International Travel Clinic; Main Operation Room – Post-Acute Care Unit; Nuclear Lab; Nutrition Clinic; Pathology; Anthology; Point of Care Testing; Respiratory Care Therapist; Respiratory Therapist; Social Work; Special Education; Student Health Services; Surgery; Clinical Toxicology and Therapeutic Drug Monitoring; Transplant Center Admin; Transplant/General Surgery Unit; Vanderbilt Medical Group Claims Management; Vanderbilt Medical Group Expected Reimbursement; Vanderbilt University Non-Employee Payments; Walk-in Clinics–Hwy 96; Office of Research	Center for Human Nutrition; Vanderbilt Multiple Sclerosis Center; Department of Pharmacology; Family Resource Center; Heart Institute; Heart Station-EKG; Hospital Guest Services; Informatics Center; Kennedy Center; Medicine – Rheumatology; Mental Health Center; Myelosuppression; Neuro – Epilepsy Lab; Palliative Care/Inpatient Med; School of Nursing Control Center; Psychiatry and Neurology; Radiology Administration; Rehab; Student Workers; Rheumatology Infusion; School of Nursing Academic Faculty Other; School of Nursing Research Admin; Sleep Lab; Specimen Receiving; Spring Hill Walk-in Clinic; Vanderbilt Medical Group Executive Leadership; Vanderbilt Orthopaedic Institute – Rehab Services; Vanderbilt Psychiatric Hospital Social Work; Vanderbilt University Police Department

members than employees in the other 2 patterns. Second, the interaction pattern for P_1 has the highest graph density and cluster coefficient, which demonstrates that there is more collaboration among team members than in the other patterns. Third, the interaction pattern for P_1 has the shortest average path length, which indicates that a pair of members in the interaction pattern tend to have a more direct line of communication than in the other 2 patterns.

Beyond the differences in network characteristics between patient groups, there are also differences in operational areas between these groups. Table 2 depicts the operational areas that were observed in only one interaction pattern per each pair of patterns. It can be seen that P_1 seems to be associated more with nursing technologies and executive leadership, while P_2 has a greater affinity to human genetics, the nuclear lab, clinical pharmacology, and transplant, and P_3 is more related to heart and mental health–related operations. Although there are variations in operational areas between these groups, the overall differences are much smaller.

Relationships between interaction patterns and LOS

The relationships between interaction patterns and LOS are depicted in Figure 3. It can be seen that P_1 achieves the shortest LOS, being 0.61 days shorter than P_2 and 0.43 days shorter than P_3. The differences in LOS between P_1 and P_2, and between P_1 and P_3, were found to be significant at the 95% confidence level. There were no significant differences in LOS detected between P_2 and P_3.

Similarities of confounding factors between patient groups

The differences between the patient groups in terms of PheWAS codes, CPT codes, insurance program, age, number of encounters to date, admission month, and access action types are depicted in Table 3. It can be seen that each encounter group exhibits similar distributions in terms of (1) PheWAS codes (PCC > 0.97, $P < 2.62 \times 10^{-26}$); (2) CPT codes (PCC > 0.98, $P < 1.67 \times 10^{-26}$); (3) insurance program (PCC > 0.98, $P < 1.73 \times 10^{-5}$); (4) age (PCC > 0.88, $P < 3.36 \times 10^{-12}$); (5) number of encounters to date (PCC > 0.89, $P < 2.14 \times 10^{-11}$); and (6) access action types (PCC > 0.98, $P < 7.15 \times 10^{-58}$). Admission month was found to be dissimilar between encounter groups ($P > 0.05$). As shown in Table 1, most of the patients in P_1 were admitted to the hospital in May, June, and July, whereas most patients in P_2 and P_3 were admitted in September, October, November, and December. Although the 3 interaction patterns dealt with different number of patients in each admission month, our results suggest that they provided care for similar patients (eg, PheWAS codes, CPT codes, age, number of encounters to date, access action types, and insurance program).

Although the overall differences in PheWAS codes, CPT codes, and access action types between patient groups were not significant, there were still small differences in several specific codes or action types worth noting. To provide insight into these codes and action types, Table 4 depicts the top 10 PheWAS codes, CPT codes, and access action types that exhibited the greatest differences between patient groups. It can be seen that P_2 has: (1) 5.6% and 1.1% more patients associated with "350.2: abnormality of gait" than P_1 and P_3, respectively; (2) >8.7% and >2.4% more patients associated with "82435: blood chloride," "82310: calcium," "84520: urea nitrogen," "82565: creatinine" than P_1 and P_3, respectively; and (3) >0.19% and >0.19% more patients associated with "Immunizations," "Emergence," "Discharge Letter," "Clinic Note," "Echo," and "Colonoscopy Operative Report" than P_1 and P_3, respectively.

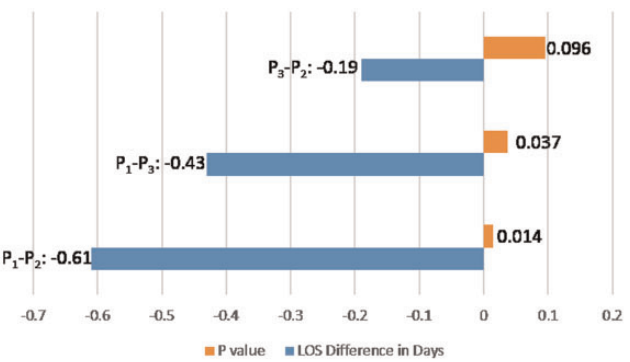

Figure 3. The difference in length of stay (LOS) for each pair of patient encounter groups. Inferred patient encounter groups are represented as P1 ($n = 428$), P2 ($n = 1353$), and P3 ($n = 3807$).

DISCUSSION

To the best of our knowledge, this is the first study to use EMR data to study the relationship between interaction patterns and inpatient hospital LOS among resource-intensive trauma patients. In doing so, it fills a gap in the knowledge about the relationship between interaction networks of health care employees and patient outcomes in trauma setting. This study specifically found that LOS for trauma inpatients with similar distributions in age, illness, procedure burden, number of encounters to date, number of access action types, and insurance type, were associated with 3 managed interaction patterns. The finding provides evidence that in trauma care, a highly collaborative pattern of interactions (eg, large weighted degree, graph density, cluster coefficient, and short path length) is associated with a shorter LOS, which can potentially assist health care organizations (HCOs) in refining management strategies to improve interaction efficiency among health care employees to reduce LOS.

We believe that this investigation has notable implications with respect to the efficiency of resource allocation and EMR system utilization for a major hospital with a trauma center. For instance, it was found that more patients were admitted in May, June, and July (ie, patients in group 1), which suggests that HCOs might consider allocating more clinical staff who are involved in these months to potentially reduce LOS (eg, via a reduction in wait time). From the perspective of EMR system utilization efficiency, we found that patient group 1 (affiliated with the shortest LOS) had more operative notes and respiratory care notes than the other patient groups, which suggests that HCOs could consider encouraging their employees to utilize EMR systems to add more informative evidence

Table 3. A similarity analysis of the factors potentially confounding the relationship between length of stay and interaction patterns.

Patient groups		P_1 vs P_2	P_1 vs P_3	P_2 vs P_3
PheWAS codes	PCC	0.9791	0.9866	0.9929
	P-value	2.62×10^{-26}	1.31×10^{-27}	4.32×10^{-29}
CPT codes	PCC	0.9855	0.9934	0.9929
	P-value	1.67×10^{-26}	2.56×10^{-29}	3.87×10^{-29}
Insurance programs	PCC	0.9808	0.9938	0.9810
	P-value	1.73×10^{-5}	5.98×10^{-7}	1.68×10^{-5}
Age	PCC	0.8858	0.9479	0.9644
	P-value	3.36×10^{-12}	1.86×10^{-17}	4.82×10^{-20}
Number of health care encounters to date	PCC	0.8963	0.9471	0.9569
	P-value	2.14×10^{-11}	2.39×10^{-15}	1.43×10^{-16}
Admission month	PCC	−0.3089	−0.1598	0.5173
	P-value	0.33	0.62	0.08
Access action types	PCC	0.9888	0.9921	0.9987
	P-value	7.15×10^{-58}	4.63×10^{-63}	5.38×10^{-90}

Inferred patient encounter groups are represented by P_1 (428 patient encounters), P_2 (1353), and P_3 (3807). *Abbreviations*: PheWAS: phenome-wide association study, which grouped ICD-9 codes together; PCC: Pearson correlation coefficient.

Table 4. The largest difference in terms of PheWAS codes, clinical procedure terminology codes, and access action types between the three patient groups

Patient groups	[% Patients in P_1 and P_2] (difference)	[% Patients in P_1 and P_3] (difference)	[% Patients in P_2 and P_3] (difference)
PheWAS code (%)			
1009: Injury, Not otherwise specified	[46.7, 37.7] (9)	[46.7, 42.3] (4.3)	[37.7, 42.3] (−4.6)
509: Respiratory failure; insufficiency; arrest	[22.7, 13.8] (8.9)	[22.7, 20] (2.7)	[13.8, 20] (−6.2)
1008: Internal injury to organs	[43.4, 35] (7.6)	[43.5, 39] (4.5)	[35, 39] (−4)
338: Pain, not elsewhere classified	[25, 17] (8)	[25, 22] (3)	[17, 22] (−5)
807: Fracture of ribs	[40.1, 32.5] (7.6)	[40.1, 33.5] (6.6)	[32.5, 33.5] (−1)
871: Open wounds of extremities	[13, 7] (6)	[13, 9] (4)	[7, 9] (−2)
508: Pulmonary collapse; interstitial/compensatory emphysema	[41, 35] (6)	[41, 38] (3)	[35, 38] (−3)
250.42: Other abnormal glucose	[19.6, 13.9] (5.7)	[19.6, 16.7] (2.9)	[13.9, 16.7] (−2.8)
350.2: Abnormality of gait	[4.1, 9.7] (−5.6)	[4.1, 9.6] (−5.5)	[9.7, 9.6] (1.1)
507: Pleurisy; pleural effusion	[24.2, 18.7] (5.5)	[24.2, 21.2] (3)	[18.7, 21.2] (−2.5)
CPTs			
82 435: Pathology and Laboratory, Assay of blood chloride	[2.5, 15.3] (−12.8)	[2.5, 9.3] (−6.8)	[15.3, 9.3] (6)
99 232: Other Medical Services, Subsequent hospital care	[56.1, 43.4] (12.7)	[56.1, 48.7] (7.4)	[43.4, 48.7] (−5.3)
82 374: Pathology and Laboratory, Assay, blood carbon dioxide	[2.5, 14.5] (−12)	[2.5, 8.9] (−6.4)	[14.5, 8.9] (−5.6)
82 310: Pathology and Laboratory, Assay of calcium	[3.5, 14.5] (−11)	[3.5, 9.4] (−5.9)	[14.5, 9.4] (5.1)
90 732: Other Medical Services, Pneumococcal vaccine	[7.8, 17.6] (−9.8)	[7.8, 17.3] (−9.5)	[17.6, 17.3] (0.3)
84 520: Pathology and Laboratory, Assay of urea nitrogen	[14.3, 24.2] (−9.9)	[14.3, 21.5] (−7.2)	[24.2, 21.5] (2.7)
94 003: Other Medical Services, Vent management inpatient, subsequent day	[26.1, 16.5] (9.6)	[26.1, 22] (4.1)	[16.5, 22] (−5.5)
71 010: Radiology, Chest X-ray	[57.6, 48.4] (9.2)	[57.6, 54.2] (3.4)	[48.4, 54.2] (−5.8)
99 253: Other Medical Services, Inpatient Consultation	[27.3, 18.1] (9.2)	[27.3, 24.3] (3)	[18.1, 24.3] (−6.2)
82 565: Pathology and Laboratory, Assay of creatinine	[15.5, 24.2] (−8.7)	[15.5, 22] (−6.5)	[24.4, 22] (2.4)
Access action types			
Discharge note	[3.4, 1.6] (1.8)	[3.4, 0.38] (3.05)	[1.6, 0.38] (1.22)
Immunizations	[1.21, 2.18] (−0.97)	[1.21, 0.14] (1.07)	[2.18, 0.14] (2.04)
Emergence	[0.24, 0.72] (−0.48)	[0.24, 0.02] (0.22)	[0.72, 0.02] (0.70)
Discharge letter	[1.32, 1.77] (−0.45)	[1.32, 0.15] (1.17)	[1.77, 0.15] (1.62)
Clinic note	[3.12, 3.38] (−0.26)	[3.12, 0.35] (2.77)	[3.38, 0.35] (3.03)
Operative note	[3.35, 3.09] (0.26)	[3.35, 0.38] (2.97)	[3.09, 0.38] (2.71)
Echo	[0.46, 0.71] (−0.25)	[0.46, 0.05] (0.41)	[0.71, 0.05] (0.66)
Admission note	[1.01, 0.77] (0.24)	[1.01, 0.11] (0.90)	[0.77, 0.11] (0.66)

(continued)

Table 4. continued

Patient groups	[% Patients in P_1 and P_2] (difference)	[% Patients in P_1 and P_3] (difference)	[% Patients in P_2 and P_3] (difference)
Respiratory care	[4.03, 3.83] (**1.2**)	[4.03, 0.45] (**3.58**)	[3.83, 0.45] (**3.38**)
Colonoscopy operative report	[0, 0.19] (−**0.19**)	[0, 0] (0)	[0.19, 0] (**0.19**)

(eg, operative notes) to improve their communication quality. This, in turn, could lead to greater efficiencies and potentially reduce the time spent on interpreting information required for care.

While this investigation indicates that data-driven methods can provide insight into the degree to which interaction patterns are associated with LOS, there are several limitations that should be recognized, which can serve as guidance for future investigations.

First, the findings suggest that more collaboration between health care employees (through an EMR system) is associated with better outcomes; however, the reason why is unclear. In particular, our investigation focused on the statistical association and not the semantic aspects of why collaboration occurs. For instance, it is not evident why the 3 interaction patterns exhibit different network structures if they provide care for similar patient populations (ie, similar distributions in terms of PheWAS code, CPT code, insurance type, number of encounters to date, number of access action types, and age). Although we observed differences in admission months between patient groups, the association between the interaction patterns and admission seasons is unclear. As such, it is worthwhile to investigate additional factors (eg, admission season, specific traumatic injury, and historical medication utilization) that may be influencing these 3 dominant patterns of collaboration.

Second, this study investigated the interaction patterns of health care employees via indirect interactions (as documented by EMR systems), but neglected direct communication in the physical world. Although the 67 different types of access actions include almost every aspect of interactions in the EMR system (eg, historical and physical documents, problem lists, respiratory care reports, clinical communication, operative reports, and progress reports), there may be discordance between the interactions that manifest in face-to-face situations and those that happen in the EMR system. The criteria of meaningful use for EMR systems have existed for a number of years (and meaningful use is now its third stage), there are still interactions in the physical clinical world that are not documented in EMR systems.[49] This missing information may influence the interaction patterns and patient encounter groups we inferred from the EMR data.

Third, further investigation is needed to uncover the causal factors behind the differences in interaction patterns. There are, in fact, many potential factors that influence interactions, such as the season a patient is admitted, the specific injury sustained, the socioeconomic status of the patient, and the number of available beds in the next receiving facility. At the same time, the temporal relationship between health care employees may also play a role in patient outcomes,[22,29] and thus is ripe for further investigation.

Fourth, beyond LOS, additional care quality measurements (eg, survival rate, days in the intensive care unit, mortality) may need to be included to assess their relationship with interaction patterns.

Fifth, this is a data-driven study, which is different from a traditional hypothesis-driven study, that was restricted to a strict trauma population definition (eg, intervention-based study and clinical trials), and thus further investigation is required to interpret evidence we learned from the data and translate it into clinical practice.

Sixth, this investigation was based on data from a single academic medical center. Replication of this study using data from other health care organizations is necessary to confirm these findings.

CONCLUSIONS

This study leveraged data-driven methodologies to infer interaction patterns from EMR utilization and illustrate their association with LOS for trauma patients. This study specifically shows that interaction patterns with a high level of collaboration are associated with shorter hospital LOS for trauma patients. This finding is notable because it provides evidence for HCOs to do further investigations to determine causal factors leading to the differences in interaction patterns and, subsequently, differences in LOS.

COMPETING INTERESTS

The authors have no competing interests to declare.

FUNDING

This research was supported, in part, by the National Institutes of Health under grants R00LM011933, R01LM010685, R01GM120484, and K23LH125670.

CONTRIBUTORS

YC performed data collection and analysis, methods design, hypotheses design, experiment design, and evaluation and interpretation of the experiments, and wrote the manuscript. CM and MP performed hypotheses design, and interpretation of experiments, and wrote the manuscript. BM performed hypotheses design and evaluation and interpretation of the experiments, and wrote the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

REFERENCES

1. US Centers for Medicare and Medicaid Service. *National Health Expenditures 2016 Highlights*. 2017; https://www.cms.gov.
2. Hadley J. Insurance coverage, medical care use, and short-term health changes following an unintentional injury or the onset of a chronic condition. *JAMA*. 2007;297(10):1073–84.

3. May JH, Cunningham PJ. Tough trade-offs: medical bills, family finances and access to care. *Issue Brief Cent Stud Health Syst Change*. 2004;85:1–4.
4. Rice D, MacKenzie E. *Cost of Injury in the United States: a Report to Congress*. San Francisco, CA: Institute for Health & Aging, University of California and Injury Prevention Center, Johns Hopkins University; 1989.
5. Alosh H, Li D, Riley III LH, *et al*. Health care burden of anterior cervical spine surgery: national trends in hospital charges and length of stay, 2000–2009. *Clin Spine Surgery*. 2015;28(1):5–11.
6. DeRienzo C, Kohler JA, Lada E, Meanor P, Tanaka D. Demonstrating the relationships of length of stay, cost and clinical outcomes in a simulated NICU. *J Perinatol*. 2016;36(12):1128–31.
7. Medicare Payment Advisory Commission. *Healthcare Spending and the Medicare Program*. Washington, DC: Medicare Payment Advisory Commission; 2016.
8. Collins TC, Daley J, Henderson WH, Khuri SF. Risk factors for prolonged length of stay after major elective surgery. *Ann Surg*. 1999;230(2):251–59.
9. Oster ME, Dawson AL, Batenhorst CM, Strickland MJ, Kleinbaum DG, Mahle WT. Relationship between resource utilization and length of stay following tetralogy of Fallot repair. *Congenit Heart Dis*. 2013;8(6):535–40.
10. Friedman DM, Berger DL. Improving team structure and communication: a key to hospital efficiency. *Arch Surg*. 2004;139:1194–98.
11. Suarez JI, Zaidat OO, Suri MF, *et al*. Length of stay and mortality in neurocritically ill patients: impact of a specialized neurocritical care team. *Crit Care Med*. 2004;32(11):2311–17.
12. Patel MS, Patel N, Small DS, *et al*. Change in length of stay and readmissions among hospitalized medical patients after inpatient medicine service adoption of mobile secure text messaging. *J General Int Med*. 2016;31(8):863–70.
13. Southern WN, Berger MA, Bellin EY, Hailpern SM, Arnsten JH. Hospitalist care and length-of-stay in patients requiring complex discharge planning and close clinical monitoring. *Arch Int Med*. 2007;167(17):1869–74.
14. Georgiou A, Lockey DJ. The performance and assessment of hospital trauma teams. *Scand J Trauma, Resuscitation Emerg Med*. 2010;18:66.
15. Fan M, Petrosoniak A, Pinkney S, *et al*. Study protocol for a framework analysis using video review to identify latent safety threats: trauma resuscitation using in situ simulation team training (TRUST). *BMJ Open*. 2016;6(11):e013683.
16. Burke RV, Demeter NE, Goodhue CJ, *et al*. Qualitative assessment of simulation-based training for pediatric trauma resuscitation. Surgery. 2016;161(5):1357–66.
17. Mayer DK, Deal AM, Crane JM, *et al*. Using survivorship care plans to enhance communication and cancer care coordination: results of a pilot study. *Oncol Nurs Forum*. 2016;43(5):636–45.
18. Rutledge RI, Domino ME, Hillemeier MM, Wells R. The effect of maternity care coordination services on utilization of postpartum contraceptive services. *Contraception*. 2016;94(5):541–47.
19. Kitzman P, Hudson K, Sylvia V, Feltner F, Lovins J. Care coordination for community transitions for individuals post-stroke returning to low-resource rural communities. *J Community Health*. 2016;42(3):565–72.
20. Janevic MR, Baptist AP, *et al*. Effects of pediatric asthma care coordination in underserved communities on parent perceptions of care and asthma-management confidence. *J Asthma*. 2016;30:1–6.
21. Chen Y, Ghosh J, Bejan CA, *et al*. Building bridges across electronic health record systems through inferred phenotypic topics. *J Biomed Inform*. 2015;55:482–93.
22. Yan C, Chen Y, Li B, *et al*. Learning clinical workflows to identify subgroups of heart failure patients. *AMIA Annu Symp*. 2016:1248–57.
23. He S, Gurr G, Rea S, Thornton SN. Characterizing the structure of a patient's care team through electronic encounter data analysis. *Stud Health Technol Inform*. 2015;216:21–25.
24. Coffey C, Wurster LA, Groner J, *et al*. A comparison of paper documentation to electronic documentation for trauma resuscitations at a level I pediatric trauma center. *J Emerg Nurs*. 2015;41:52–56.
25. Huyvetter C, Lang AM, Heimer DM, Cogbill TH. Efficiencies gained by using electronic medical record and reports in trauma documentation. *J Trauma Nurs*. 2014;21(2):68–71.
26. Sollie A, Sijmons RH, Helsper C, Numans ME. Reusability of coded data in the primary care electronic medical record: a dynamic cohort study concerning cancer diagnoses. *Int J Med Inform*. 2017;99:45–52.
27. McGrath K, Foster K, Doggett P, *et al*. EMR-based intervention improves lead screening at an urban family medicine practice. *Fam Med*. 2016;48(10):801–04.
28. Chen Y, Lorenzi NM, Sandberg WS, Wolgast K, Malin BA. Identifying collaborative care teams through electronic medical record utilization patterns. *J Am Med Inform Assoc*. 2017;24:e1:e111–20.
29. Chen Y, Xie W, Gunter C, *et al*. Inferring clinical workflow efficiency via electronic medical record utilization. *AMIA Annu Symp*. 2015:416–25.
30. Gao C, Kho AN, Ivory C, Osmundson S, Malin B, Chen Y. Predicting Length of Stay for Obstetric Patients via Electronic Medical Records. *Stud Health Technol Inform*. 2017;245:1019–23.
31. Collins N, Miller R, Kapu A, *et al*. Outcomes of adding acute care nurse practitioners to a Level I trauma service with the goal of decreased length of stay and improved physician and nursing satisfaction. *J Trauma Acute Care Surg*. 2014;76(2):353–57.
32. Giuse DA. Supporting communication in an integrated patient record system. *AMIA Annu Symp Proc*. 2003:1065.
33. Denny JC, Ritchie MD, Basford MA, *et al*. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*. 2010;26:1205–10.
34. Denny JC, Bastarache L, Ritchie MD, *et al*. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol*. 2013;31:1102–11.
35. Chen Y, Lorenzi N, Nyemba S, Schildcrout JS, Malin B. We work with them? Health workers interpretation of organizational relations mined from electronic health records. *Int J Med Inform*. 2014;83:495–506.
36. Van Mechelen I, Bock HH, De Boeck P. Two-mode clustering methods: a structured overview. *Stat Methods Med Res*. 2004;13(5):363–94.
37. Huang S, Wang H, Li D, Li T. Spectral co-clustering ensemble. *Knowledge-Based Sys*. 2015;84:46–55.
38. Kluger Y, Basri R, Chang JT, Gerstein M. Spectral biclustering of microarray data: coclustering genes and conditions. *Genome Res*. 2003;13(4):703–16.
39. NCEPOD. Trauma: Who Cares?: a Report of the National Confidential Enquiry Into Patient Outcome and Death (2007). London, UK: NCEPOD; 2007.
40. Nguyen HV, Bai L. Cosine similarity metric learning for face verification. *Proceedings of the Asian Conference on Computer Vision*. Springer: Berlin; 2010:709–20.
41. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proc AAAI In Conf Weblogs and Social Media*. 2009;8:361–62.
42. Boccaletti S, Latora V, Moreno Y, *et al*. Complex networks: Structure and dynamics. *Phys Rep*. 2006;424(4):175–308.
43. Blondel V, Guillaume J, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stats Mechan*. 2008;2008(10):P1000.
44. McCullagh P. Generalized linear models. *Eur J Operational Res*. 1984;16(3):285–92.
45. Austin PC, Rothwell DM, Tu JV. A comparison of statistical modeling strategies for analyzing length of stay after CABG surgery. *Health Services Outcomes Res Methodol*. 2002;3(2):107–33.
46. David C, Hoaglinab E. The hat matrix in regression and ANOVA. *Am Stat*. 1978;32(1):17–22.
47. Fisher RA. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*. 1915;10(4):507–21.
48. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proc AAAI International Conference on Web and Social Media*. 2009;8:361–2.
49. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. *N Engl J Med*. 2010;363:501–04.