Inferring EHR Utilization Workflows through Audit Logs

Xinmeng Zhang, BS[1,*], Yuying Zhao, MS[1,*], Chao Yan, PhD[2],
Tyler Derr, PhD[1], You Chen, PhD[1,2]
[1]Vanderbilt University, Nashville, TN; [2]Vanderbilt University Medical Center,
Nashville, TN.

Abstract

Electronic health records (EHRs) usage and clinical workflows are intrinsically linked. To accommodate the complex care settings (e.g., emergency departments), EHR utilization workflows dynamically change in clinical practice, which in turn shapes the clinical workflows. Learning EHR workflows would provide an opportunity for healthcare organizations to enhance clinical workflows in the context of EHRs. However, very few studies investigated EHR utilization workflows executed in clinical practice. We develop a network analysis framework and apply it to EHR audit logs to infer EHR workflows. We then measure the differences in the workflows between patient subgroups divided by races via differential network analysis. We apply our framework to trauma patients admitted to the emergency department, which is one of the clinical settings that need timely support from EHR utilizations. Our results show five core EHR workflows related to Narrator, Navigator, SmartTools, Chart Review, and ED workup activities in the ED. We find EHR workflows involving Narrator, SmartTools, and BPA are different when comparing patient subgroups.

Introduction

Over the past decade, electronic health records (EHRs) have been widely deployed across the United States to advance health information interoperability and accessibility[1]. In the meantime, much evidence demonstrated their positive impact in quickening diagnosis, reducing medical errors, and supporting better patient outcomes. This is because the readily available and accurate data stored in EHRs streamlines the time-intensive processes and contributes to easier clinical decision-making[2-4]. As such, EHR utilization has been indivisibly integrated into care routines. As the major component of clinical practice, clinical workflows, which consist of a variety of tasks performed by healthcare professionals within and between care settings to deliver care, dictates EHR functions and utilizations[5,6]; while EHRs, as the information hub of the clinical practice, can, in turn, shape clinical workflows and have the potential to enhance care efficiency[7,8]. In other words, EHR usage and clinical workflows are intrinsically linked and impact each other.

EHR utilization workflows (or EHR workflows) usually start from a set of predefined or recommended processes, which are often idealistic, linear, and less likely to accommodate complex healthcare environments[7,9]. In fact, they are subject to dynamic change and evolvement given the diversity brought by multiple important factors, such as patient urgency and severeness, whether there are regional or global disease outbreaks (e.g., COVID-19 pandemic), care routine upgrades, and insurance policy changes[9-11]. Understanding how EHRs are utilized in practice can not only lighten the dark area in the digital world by uncovering important patterns that are hidden to EHR designers but also provide evidence to inspire opportunities for the improvement of their mediated clinical workflows.

EHR audit logs, which capture all of the detailed events performed by EHR users to patients' records, are one of the perfect data sources for investigating EHR utilization workflows. Notably, this type of data has been widely collected by EHR systems and utilized by researchers to support significant learning objectives, including but not limit to health process modeling, time-motion study of EHR usages of physicians, and clinical burnout. Process mining, which combines data mining and business process management techniques, has been leveraged to analyze EHR audit logs to discover EHR workflows and measure their characteristics (e.g., the duration to complete vital signs charting) in various care settings (e.g., primary, pediatric, or ophthalmic care)[12-22]. Process mining has yielded many interesting findings such as attending physician, nurse, or resident's workflows to complete EHR tasks[13,14]. However, this technique provides little information on the structures of the workflow, for instance, which are central events in a workflow and what are relationships between the central events and other events.

This paper aims to explore applying network analysis techniques (which can successfully capture complex relationships between events and provide excellent visualizations) to the EHR audit logs data to uncover the topological patterns of EHR utilization workflows. Specifically, we develop a network analysis-based framework to

---

* Equal Contribution

learn EHR utilization workflows through community detection algorithms and then measure the workflow structures via sociometric factors. Further, we investigate the disparities of EHR utilization workflows between patient racial subgroups, which is motivated by findings in previous studies that racial and ethnic biases exist in many EHR components, such as laboratory tests and clinical notes[23,24]. In this study, we focus on the trauma patients admitted to the emergency department (ED), which correspond to one of the most fast-paced and demanding clinical settings[8,25,26] that requires the efficient and accurate utilization of EHRs to support fast diagnostics and care delivery.

Materials and Methodologies

Patients and Audit logs

This study focuses on trauma patient care in the ED and uses a two-year cohort (2018-2019) of a trauma registry repository, which was fully anonymized and routinely collected from Vanderbilt University Medical Center (VUMC) between January 2017 and May 2021. The EHR audit logs (which characterize the EHR utilization details of EHR users, who provide healthcare service to patients) for each included patient in the period between their ED admission and ED discharge were collected from the Epic EHR system deployed at VUMC. Specifically, EHR utilization on each patient is logged as a sequence of time-stamped events (i.e., *user-EHR interactions*), each corresponding to a record containing four domains: user ID, patient ID, timestamp, and user-EHR interaction type (Figure 1A). The last domain is designed to describe the semantics of user-EHR interactions made by EHR users, including hundreds of types under the categories of *view*, *modify*, *export*, and *system*[27].

Summary statistics of patient demographics and the patient-level user-EHR interactions (and their types) extracted from the EHR audit logs are depicted in Table 1. A total of 3,946 patients were included in this study (1446 female [36.6%]; median age 56.1 [IQR: 35.1-72.7]; 3,332 White patients [84.4%], 452 Black patients [11.4%], 33 Asian patients [0.8%], and 162 patients with other races [4.1%]). Accordingly, we collected ~2,306,000 user-EHR interactions, which correspond to 677 unique types.
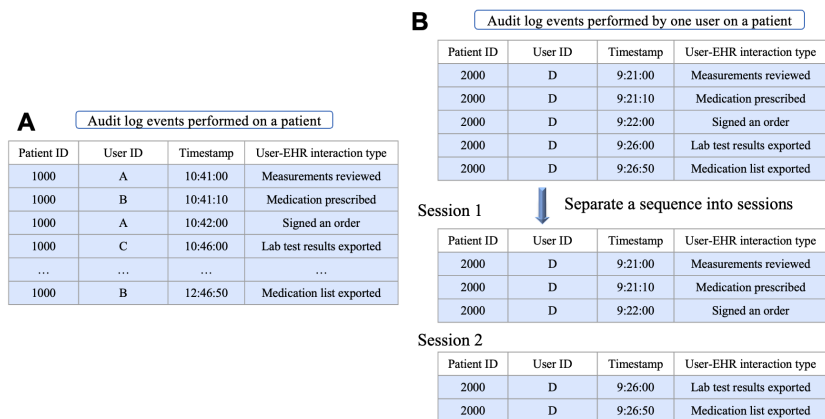

Figure 1. Two examples to illustrate **(A)** a user-EHR interaction sequence on a single patient and **(B)** the process of creating sessions from a sequence of user-EHR interactions performed by one user to a patient's EHR, respectively.

Transform sequences of user-EHR interactions into sessions

Clinical practitioners utilize EHRs to support the completion of clinical tasks. The EHR systems are highly modularized in design and the functions or components associated with the same clinical task are clustered such that clinicians can finish a sequence of user-EHR interactions within a continuous period of time that is relatively short, leading to a temporal EHR usage pattern–the alternation of dense-event period and idle period. This phenomenon naturally separates different semantics in care delivery and has been validated in the literature[13]. We first organize the sequence of user-EHR interactions of each patient at the user level (leading to multiple subsequences) and then split each subsequence into multiple sessions (Figure 1B). We assume that each subsequence represents a specific EHR task and such a sessionized process can help discover the EHR task-oriented workflows (weakening relationships between events belonging to two sessions). Following our previous findings in the same insititution[13], we set a threshold of 2 minutes between two consecutive events as a cutoff criterion to extract sessions from the sequence of user-EHR interactions.

Table 1: Summary of the dataset used in this study. a, b, c represents the first quartile, median, and third quartile. $x \pm y$ represents the mean and one standard deviation. $x\%\ y$ represents that the percentage of y patients (in a given category) is $x\%$ among all patients.

Characteristic	Distribution			
Race				
White		84.4%	3,332	
Black		11.4%	452	
Asian		0.8%	33	
Others		4.1%	162	
Age at admission		35.1, 56.1, 72.7	54.5 ± 21.9	
	White patient		**Non-White patient**	
Age at admission	38.8, 58.7, 74.7	56.6 ± 21.6	27.2, 40.6, 60.0	44.7 ± 20.5
Gender				
Male	61.7%	2,055	72.5%	445
Female	38.3%	1,277	27.5%	169
ED duration(hour)	2.2, 4.3, 6.6	5.0 ± 4.4	1.4, 3.8, 6.1	4.4 ± 3.8
Local Injury Severity Score (ISS)	10, 14, 18	15.4 ± 7.0	9, 13, 18	15.2 ± 7.8
User-EHR interactions (in ED visit-level)				
Number of interactions	237, 534, 838	597.0 ± 475.4	150, 459, 767	516.7 ± 448.2
Number of unique interactions	55, 82, 105	78.8 ± 38.7	43, 79, 104	74.3 ± 40.8

Create a directed network from sessions

Next, we build directed networks based on the extracted sessions from the study cohort. Each node in the network represents a semantic type of user-EHR interactions. If two user-EHR interactions with different types are consecutive in the same session (i.e., "A→B"), the corresponding nodes are connected with the edge direction indicating the order of appearing. The pattern of "A→A" is excluded in network construction to avoid self-looping in the directed network. The weight of an edge is defined to be the count of the two user-EHR interaction types consecutively appearing in all sessions. Following this network construction pipeline, we build one weighted directed network describing the relationships of user-EHR interaction types across all sessions and ED trauma patients performed by EHR users. Figure 2 illustrates the process of constructing a single directed network from multiple sessions.

To investigate whether there exist disparities in EHR utilization patterns between patient subgroups, we select patient race as the grouping factor and apply the aforementioned pipeline to create two sub-networks–one based on the White patients (race majority), and the other based on the Non-White patients (race minority).
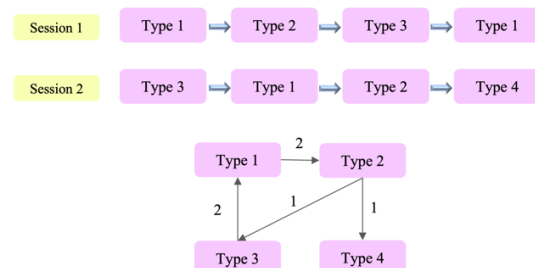

Figure 2. Illustration of the process of creating a directed network of user-EHR interaction types from sessions.

Identify cores and bottlenecks from the networks

We propose to use three sociometric to characterize networks. We use degree and PageRank score[28] to quantify the local and global centrality of a node in the network, respectively. Meanwhile, betweenness centrality is leveraged to measure the efficiency and effectiveness of a node connecting pairs of other nodes in the network. The degree centrality indicates the strength of a node to its one nearest neighbor, which characterizes the local influence of this node. PageRank score characterizes the strength of a node connected to other nodes that are also strongly connected to others (hubs connected to other hubs) across the whole network. A user-EHR interaction type with a high PageRank score is likely to be a core event in an EHR workflow. The betweenness centrality characterizes the number of shortest

paths between two other nodes that pass through a specific node. A node with a high betweenness centrality, i.e., the bottlenecks of the network, enables more efficient connections/interactions (shortest path) in EHR workflows.

For each of the three networks, Whole, White, and Non-White, we use *NetworkX*[29], an open-source network analysis tool to calculate the degree, PageRank, and betweenness centrality of each node. Further, we investigate whether differences in degree, PageRank, and betweenness centrality of the networks between White and Non-White patients are statistically significant. The degree is normalized by the number of patients in the corresponding patient group. We use a Mann-Whitney U test with a significant level of 0.05 since sociometric factors are not Gaussian distribution.

Infer EHR workflows through community detections

To detect high-quality communities from the constructed network, we filter weak relationships by the following steps. We first normalize the edge weights to [0, 1] by dividing the maximum weight in the network and then filter out those whose weights are less than a predefined threshold τ. After the adjustment, the size and complexity of network G are reduced, whereas more significant nodes and their relationships are maintained. In this study, we focus on the largest connected component C_l^G in each network because they dominate the corresponding networks in our settings. We select a threshold τ from a continuous range that can yield a stable C_l^G for all of the three networks by detecting the continuous range candidates of τ (if any) such that in each range changing the value of τ does not change the size of C_l^G.

We then perform the Louvain algorithm on C_l^G of each network to obtain the clusters and visualize them via Gephi[30] (open-source network analysis and visualization software). Louvain algorithm is an unsupervised community detection algorithm[31], which detects clusters from a network via iterative modularity optimization and community aggregation. We use the Louvain algorithm over other community detection algorithms, such as DBSCAN[32], as it does not require prior assumptions on the network, which could introduce implicit bias. Further analysis is conducted to align clusters learned from the three networks and measure differences in the clusters between White and Non-White patients.

Differential network analysis: compare differences in topologies between networks

The differential network analysis (DiNA)[33] focuses on the differences in the topologies between networks. Our usage of DiNA is originally motivated by its common use in Bioinformatics studies. For example, cells from healthy people and patients could be in different states. DiNA could identify key cells from two cell networks of two biological processes. In this study, we measure differences in connection structures of user-EHR interaction types in networks built from White vs Non-White trauma ED patients.

We employ two algorithms to measure network difference as introduced by Lichtblau *et al*[33]:

- Differential Degree Centrality (DDC). Assume A and B are two networks. Let $d^A(v)$ and $d^B(v)$ denote the degrees of a node v in network A and B, respectively, $d^A(max)$ and $d^B(max)$ denote the maximum degrees in network A and B, respectively. The degree of each node is first normalized by the maximum degree value in the corresponding network. DDC of a given node is defined as the absolute difference of the normalized degree of this node in network A and B.

$$DDC(v) = \left| \frac{d^A(v)}{d^A(max)} - \frac{d^B(v)}{d^B(max)} \right| \tag{1}$$

- Differential PageRank Centrality (DPC). Let $S_{PR}^A(v)$ and $S_{PR}^B(v)$ denote the PageRank score of node v in network A and B, respectively, $S_{PR}^A(max)$ and $S_{PR}^B(max)$ denote the maximum PageRank scores from network A and B, respectively. Similar to DDC, DPC of a given node is defined as the absolute difference of the normalized PageRank scores of this node in network A and B.

$$DPC(v) = \left| \frac{S_{PR}^A(v)}{S_{PR}^A(max)} - \frac{S_{PR}^B(v)}{S_{PR}^B(max)} \right| \tag{2}$$

Results

Constructed Networks: Whole, White, and Non-White

Using the 2-minute threshold as the cutoff criterion to create sessions, we obtain approximately 236K sessions with an average length of 5 events and a standard derivation (STD) of 6.1. In the constructed Whole network, there are 677

nodes with 16,319 directed edges. The average degree, PageRank, and betweenness centrality are 33.6 with an STD of 45.9, 0.001 with an STD of 0.003, and 0.001 with an STD of 0.006, respectively. We observe that "*Visit Navigator template loaded*", "*A SmartText used in SmartTools or reports*", and "*A SmartLink used in SmartTools or reports*" are the top-3 user-EHR-interaction types in all three metrics (i.e., degree, PageRank, and betweenness scores). This indicates that these user-EHR interaction types correspond to the core events in both the local and global EHR workflows, and meanwhile act as the bridge for information transfer along with EHR workflows.

There are 666 and 530 nodes, and 15,458 and 8,136 edges in the White and Non-White patients' networks, respectively. 519 nodes appear in both networks from White patients and Non-White patients. For comparing the differences between these two networks, we only analyze the sociometric on the common nodes. Table 2 summarizes the sociometric calculated from three networks. We observed that the average normalized degrees are significantly different between the White patient network and the Non-White patient network. It indicates that there are fewer user-interaction type transitions per patient for White patients than Non-White patients.

Table 2. Summary of the sociometric on the whole network and the sub-networks. $x \pm y$ represents the mean and one standard deviation. P-value denotes the two-sided Mann-Whitney U test with a significant level of 0.05 comparing the sociometric between the White and Non-White networks.

Sociometric	Distribution			
	Whole network	White patient network	Non-White patient network	P-value
Degree	33.6 ± 45.9	0.012 ± 0.013	0.036 ± 0.049	< 0.001
PageRank	0.001 ± 0.003	0.001 ± 0.004	0.001 ± 0.004	0.128
Betweenness	0.001 ± 0.006	0.002 ± 0.007	0.002 ± 0.009	0.177

EHR workflows

Figure 3 shows the size of C_l^G as a function of the threshold τ, where we observe three plateaus in terms of C_l^G size. The middle points of the three areas in the rectangles are selected as candidate values of τ. Since using the two larger thresholds will lead to an excessive loss of information, we use $\tau = 0.08$, which results in 24 nodes in the Whole network and the White network, and 25 nodes in the Non-White network, respectively.
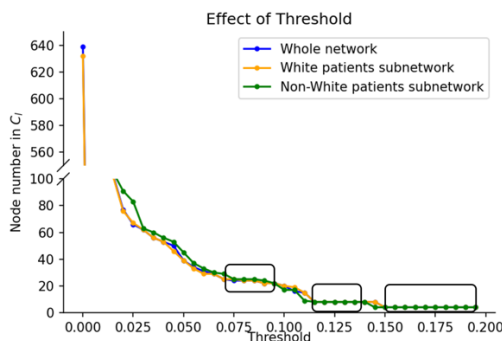

Figure 3. The effect of threshold values on the size of C_l^G.

By applying the Louvain algorithm, we detected five clusters as shown in Figure 4 for the Whole network.

Users' interactions with ED Narrator. Users can use the ED Narrator to jump from toolbox to toolbox, documenting assessments or completing tasks as needed. The Narrators are designed to be used in any order. There are three sections in the Narrator: the toolbox on the left and right and the event log in the center. "*Event log was loaded in Narrator*" has the highest local influence in this cluster, while also being a core event in the overall EHR workflows.

Users' interactions with ED Navigator. ED Navigators are designed to match the order of a workflow and users will typically work from the top section to the bottom section. "*Visit Navigator template loaded*" has the most connections with other user-interaction types in this cluster. This type is also one of the top-3 user-EHR-interaction types in degree, PageRank, and betweenness centrality, indicating its local and global impact and its role as a bridge in the workflows.

Users' interactions with SmartTools in handoff data. SmartTools include SmartTexts, which are templates or blocks of text, SmartPhrases, which are long words, phrases, or paragraphs, SmartLinks, which pull information from another part of the chart, and SmartLists, which are a predefined list of choices. SmartLinks help ED professionals write notes quickly by pulling or linking information from the patient's chart directly into the notes. SmartTexts are standard templates or blocks of text used to write notes for ED visits or problems a healthcare professional treats. There are strong connections between "*A SmartText used in SmartTools or reports*", and "*A SmartLink used in SmartTools or reports*", indicating these two user-interaction types are often utilized sequentially.

Chart review activities, including reviewing notes, encounters, orders, diagnoses, rooming plans, *etc*. In Chart Review Activity, ED professionals can find patient health information that's been documented in the EHR from past hospital encounters as well as the current hospitalization. This cluster does not have strong connections with user-interaction types from other clusters.

ED workup activities. The workup activity has results, vital signs, and the ED course. For instance, when results (labs or vital signs) are returned, there will be an icon that identifies there are new results in the lab or vital sign widget.

The results show that there are five relatively independent EHR workflows, while those workflows are also connected through several user-interaction types. For instance, ED workup activity is connected to Narrator, Navigator, and SmartTools through SmartText, Navigator template, and Narrator event log. Chart review is connected to SmartTools through orders.
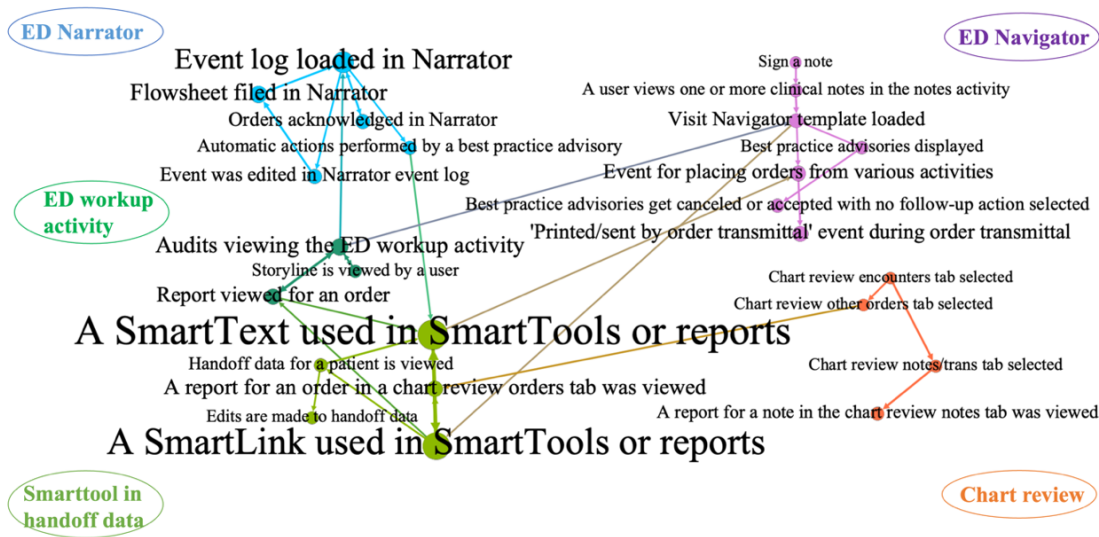

Figure 4. Five clusters detected for the Whole network. The width of the edge illustrates edge weight, and the size of a node illustrates a node's PageRank score.

EHR workflows differences

Based on the White patients' network and Non-White patients' network, we observe that nodes in the C_l^{White} are the same as the nodes in C_l^{Whole} and 24 of 25 nodes in $C_l^{Non-White}$ are the same. It indicates that these user-interaction types are the most frequently associated with all patients. The one extra node in $C_l^{Non-White}$ is "*The chart review notes report was viewed*". This user-interaction type has a similar meaning to another user-interaction type: "*A report for a note in the chart review notes tab was viewed*". We obtain five clusters from C_l^{White} and $C_l^{Non-White}$, similar to the results from C_l^{Whole} after applying the Louvain algorithm. This verifies that our algorithm finds meaningful and general workflow patterns in ED trauma patient care using user-interaction types. The Narrator cluster and the SmartTools workflow cluster are different for White patients and Non-White patients. Figure 5 shows these two clusters specifically. We observe that "*Automatic actions performed by a best practice advisory*" (BPA) is often performed with the Narrator workflow for White patients while being performed with the SmartTools workflow for Non-White patients.
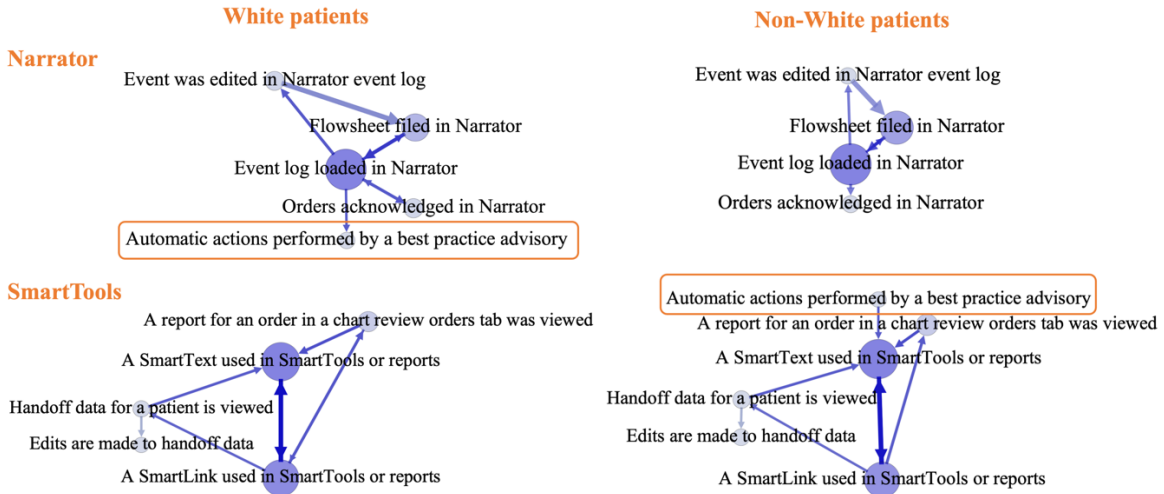
Figure 5. The Narrator and SmartTools workflow differences on patient groups. The width of the edge illustrates edge weight, and the size of a node illustrates a node's PageRank score.

Differential network analysis is conducted on the 24 intersected nodes between C_l^{White} and $C_l^{Non-White}$. From the differential degree centrality (DDC) measure between White and Non-White patients' networks, five user-interaction types have non-negative DDC values: "*Event log was loaded in Narrator*", "*A SmartLink used in SmartTools or reports*" (SL), "*Automatic actions performed by a best practice advisory*" (BPA), "*Chart review notes/transaction tab selected*", and "*Report for an order in a chart review orders tab was viewed*". Except for "*Chart review notes/transaction tab selected*", the other four user-interaction types have higher degrees in the White patients' network than those in the Non-White patients' network.

Figure 6 shows the visualization of the differential network. The user-interaction types that have the top-4 DPC values are related to the Narrator workflow. These Narrator-related user-interaction types have higher PageRank values in the Non-White patients' network. We observe that BPA has the 5th highest DPC value and SL has the 9th highest DPC value. Both user-interaction types have higher PageRank values in the White network. It is noted that SL and BPA have both higher global (higher PageRank value) and local impacts (higher degree) on the EHR workflows in the White patients' network than that in the Non-White.
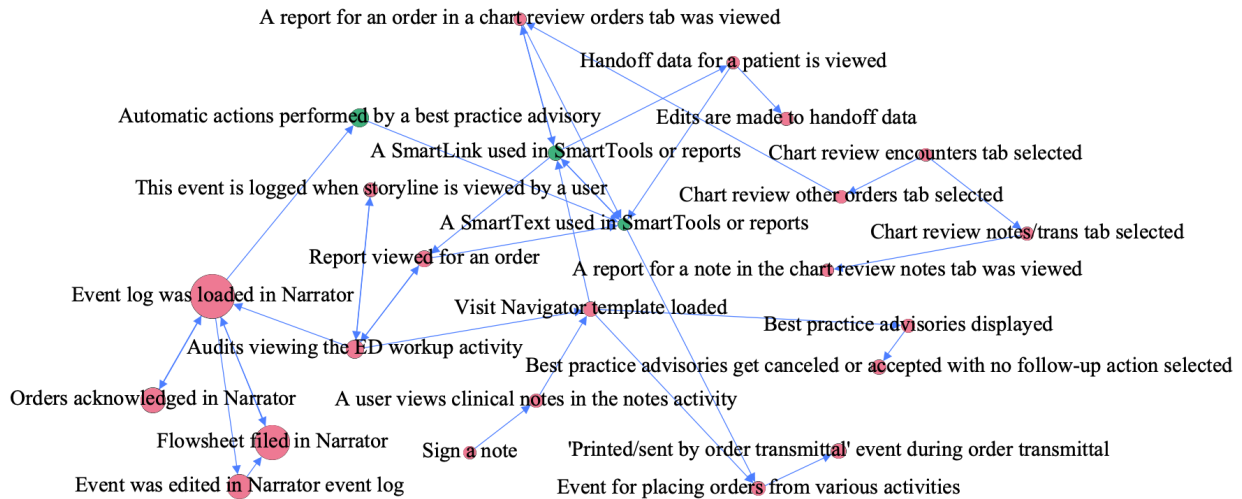

Figure 6. Differential PageRank centrality. The node size to represent the absolute difference of the values of DPC. The pink color shows that the Non-White patients' network has the higher DPC values. The green color shows that the White patients' network has higher DPC values.

Discussion

The ED is one of the most fast-paced and demanding medical specialties that admit trauma patients of various ages, diagnoses, and treatment plans. Our study's focus on ED trauma patient care is an attempt to capture EHR workflows in such a diverse and complex environment. We apply network analysis and community detection algorithm to learn EHR workflow patterns and measure differences in the content (e.g., Narrator, Navigator, SmartText, SmartLink, BPA) and structure (user-interaction types with degree, PageRank, and betweenness scores) of the workflow pattern between White and Non-White patients in systematic ways.

Network analysis results show that user-interaction types related to Narrator, Navigator, SmartText, and SmartLink have the highest values of sociometric factors (degree, PageRank, and betweenness centrality), which demonstrate their critical roles in ED EHR workflows of taking care of trauma patients. In such a fast-paced setting, ED professionals check trauma patients' health status, laboratory test results, vital signs, and medication orders frequently and quickly. Narrator, Navigator, and SmartTools provide strong support to speed up those tasks in the EHR. The analysis gives us more micro-level understandings of how EHRs are utilized in ED trauma patient care and provides an opportunity to analyze EHR efficiency. For instance, we found BPA is one of the core user-interaction types in Narrator (switches between toolboxes) and Navigator (BPA displayed and then got accepted or canceled). BPA can increase the efficiency and effectiveness of evidence-based medical practice; however, alert fatigue leads users to begin to ignore them[34]. ED users heavily rely on the Narrator to manage patients' health status and medical needs through various tools. If BPA's number is not well controlled for those tools, then ED users may be disrupted by the flags and would raise issues related to care quality and patient safety. Community detection gives us macro-level insights into the EHR. The results on the whole network show that there are five core EHR workflows related to Narrator, Navigator, SmartTools, Chart Review, and ED workup activities. Regardless of clustering on the whole network or networks divided by White and Non-White patients, the five core EHR workflows are similar, suggesting the detected workflow patterns are stable. As shown in Figure 4, the 5 core EHR workflows are connected by user-interaction types, including orders and workup activities. Orders such as medication, laboratory tests, diet, and workup activities (safety) are critical tasks for ED patients. Our analysis uncovered their relationships, which may provide evidence for systematically optimizing EHR workflows as a whole rather than refining each individually.

The differential network analysis provides a comparative perspective into EHR workflow patterns in ED trauma patient care between White and Non-White patients. SmartLink, BPA, and Narrator play different roles in the EHR workflow between patient groups. SmartLink and BPA have both higher global impact (higher PageRank scores) and local impact (higher degrees) on White patients than Non-White patients, which indicates the two user-interaction types may be frequently utilized across the five core EHR workflows for White patients. The user-interaction type associated with the Narrator has a more local impact on the White patients, while a more global impact on the Non-White patients, which indicates Narrator is utilized differently for White and Non-White patients. However, the impact of such utilization behaviors on the EHR workflows, related clinical workflows, and health outcomes is unknown. It is unclear why EHR utilization differences existed in the ED EHR workflows between White and Non-White trauma patients. Further investigation on EHR workflow through other approaches such as focus group interviews with ED professionals or observations in the ED would provide more evidence to illustrate the differences.

Our study should acknowledge several limitations. First, the audit log data has been recognized as one of the most informative resources to study EHR workflows, however, our learned EHR workflow patterns have not been systematically evaluated. A previous study[35] investigated the capability of the audit log data in capturing clinician activities in EHRs shows that the log-generated breadcrumbs encounter summary can capture all interactions documented in clinical notes, with the exception of physical exams. Based on their observations, there is a high chance that the EHR workflows can be reflected in the audit log data. Second, our study only uses trauma patients' data from VUMC ED. The results might not be generalizable for other hospitals when the patients' demographic distribution is different from that in VUMC, or they use different EHR systems such as Allscripts or Cerner. Third, EHR workflow patterns are identified from the audit log data, however, it is not connected to clinical workflow, and thus its impact on clinical workflow is unknown. Further investigations are required to align EHR workflow and clinical workflow and identify EHR system usability issues to improve care quality and patient safety. Fourth, due to the complexity of associations between EHR workflows and health outcomes and page limit, this study did not measure the relationships between EHR workflows with outcomes such as length of stay, morbidity, and mortality. Fifth, our study only compares the networks built from White and Non-White patients. Further analysis could investigate the differences between patients grouped by other factors, such as genders or categories of diseases.

Conclusion

Our study develops an informatics framework to infer EHR workflows in ED trauma patient care. We analyze the content (user-interaction types) and the structure of ED EHR workflows through community detection and network analysis. Furthermore, we conduct a differential network analysis to identify differences in the EHR workflow patterns between White and Non-White patients. Our framework detects 5 core EHR workflows and their differences in user-interaction types and corresponding values of sociometric factors between White and Non-White trauma patients. Future studies can be conducted to measure the relationships of core user-interaction types in EHR workflows with health outcomes.

Acknowledgement

References
1. Joukes E, de Keizer NF, de Bruijne MC, Abu-Hanna A, Cornet R. Impact of electronic versus paper-based recording before EHR implementation on health care professionals' perceptions of EHR use, data quality, and data reuse. Appl Clin Inform. 2019;10(2):199–209.
2. Bowman S. Impact of electronic health record systems on information integrity: quality and safety implications. Perspect Health Inf Manag. 2013;10:1c.
3. King J, Patel V, Jamoom EW, Furukawa MF. Clinical benefits of electronic health record use: national findings. Health Information Technology. 2014;49(1 Pt 2):392–404.
4. Yuan N, Dudley RA, Boscardin WJ, Lin GA. Electronic health records systems and hospital clinical performance: a study of nationwide hospital data. J Am Med Inform Assoc. 2019;26(10):999–1009.
5. Olakotan OO, Yusof MM. Evaluating the alert appropriateness of clinical decision support systems in supporting clinical workflow. J Biomed Inform. 2020;106(103453):103453.
6. Blezek DJ, Olson-Williams L, Missert A, Korfiatis P. AI integration in the clinical workflow. J Digit Imaging. 2021;34(6):1435–46.
7. Zheng K, Ratwani RM, Adler-Milstein J. Studying workflow and workarounds in electronic health record-supported work to improve health system performance. Ann Intern Med. 2020;172(11 Suppl):S116–22.
8. Denton CA, Soni HC, Kannampallil TG, Serrichio A, Shapiro JS, Traub SJ, et al. Emergency physicians' perceived influence of EHR use on clinical workflow and performance metrics. Appl Clin Inform. 2018;9(3):725–33.
9. Grando MA, Vellore V, Duncan BJ, Kaufman DR, Furniss SK, Doebbeling BN, et al. Study of EHR-mediated workflows using ethnography and process mining methods. Health Informatics J. 2021;27(2):14604582211008210.
10. Coleman C, Gotz D, Eaker S, James E, Bice T, Carson S, et al. Analysing EHR navigation patterns and digital workflows among physicians during ICU pre-rounds. Health Inf Manag. 2021;50(3):107–17.
11. Grando A, Manataki A, Furniss SK, Duncan B, Solomon A, Kaufman D, et al. Multi-method study of electronic health records workflows. AMIA Annu Symp Proc. 2018;2018:498–507.
12. Furniss SK, Burton MM, Grando A, Larson DW, Kaufman DR. Integrating process mining and cognitive analysis to study EHR workflow. AMIA Annu Symp Proc. 2016;2016:580–9.
13. Chen B, Alrifai W, Gao C, Jones B, Novak L, Lorenzi N, et al. Mining tasks and task characteristics from electronic health record audit logs with unsupervised machine learning. J Am Med Inform Assoc. 2021;28(6):1168–77.
14. Jones B, Zhang X, Malin BA, Chen Y. Learning tasks of pediatric providers from Electronic Health Record audit logs. AMIA Annu Symp Proc. 2020;2020:612–8.
15. Hribar MR, Chiang MF. Response to Letter: Secondary use of electronic health record data for clinical workflow analysis. J Am Med Inform Assoc. 2018;25(7):920.
16. Hribar MR, Biermann D, Read-Brown S, Reznick L, Lombardi L, Parikh M, et al. Clinic workflow simulations using secondary EHR data. AMIA Annu Symp Proc. 2016;2016:647–56.
17. Hribar MR, Read-Brown S, Reznick L, Chiang MF. Evaluating and improving an outpatient clinic scheduling template using secondary electronic health record data. AMIA Annu Symp Proc. 2017;2017:921–9.
18. Goldstein IH, Hribar MR, Sarah R-B, Chiang MF. Quantifying the impact of trainee providers on outpatient clinic workflow using secondary EHR data. AMIA Annu Symp Proc. 2017;2017:760–9.

19. Yan C, Zhang X, Gao C, Wilfong E, Casey J, France D, et al. Collaboration structures in COVID-19 critical care: Retrospective network analysis study. JMIR Hum Factors. 2021;8(1):e25724.
20. Chen Y, Nyemba S, Malin B. Auditing medical records accesses via healthcare interaction networks. AMIA Annu Symp Proc. 2012; 2012:93.
21. Chen Y, Adler-Milstein J, Sinsky C. Measuring and Maximizing Undivided Attention in the Context of Electronic Health Records. Applied Clinical Informatics. 2022; 10.1055/a-1892-1437.
22. Chen Y, Patel MB, McNaughton CD, Malin BA. Interaction patterns of trauma providers are associated with length of stay. J Am Med Inform Assoc. 2018;25(7):790-9.
23. Pivovarov R, Albers DJ, Sepulveda JL, Elhadad N. Identifying and mitigating biases in EHR laboratory tests. J Biomed Inform. 2014;51:24–34.
24. Hripcsak G, Knirsch C, Zhou L, Wilcox A, Melton G. Bias associated with mining electronic health records. J Biomed Discov Collab. 2011;6:48–52.
25. Patel VL, Denton CA, Soni HC, Kannampallil TG, Traub SJ, Shapiro JS. Physician workflow in two distinctive emergency departments: An observational study. Appl Clin Inform. 2021;12(1):141–52.
26. Pruitt ZM, Howe JL, Hettinger AZ, Ratwani RM. Emergency physician perceptions of electronic health record usability and safety. J Patient Saf. 2021;17(8):e983–7.
27. Zhang X, Yan C, Malin BA, Patel MB, Chen Y. Predicting next-day discharge via electronic health record access logs. J Am Med Inform Assoc. 2021;28(12):2670–80.
28. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. Comput netw ISDN syst. 1998;30(1–7):107–17.
29. NetworkX — NetworkX documentation [Internet]. Networkx.org. [cited 2022 Mar 8]. Available from: https://networkx.org/
30. The open graph viz platform [Internet]. Gephi.org. [cited 2022 Mar 8]. Available from: https://gephi.org/
31. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech. 2008;2008(10):P10008.
32. Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. AAAI Press; 1996. p. 226–31.
33. Lichtblau Y, Zimmermann K, Haldemann B, Lenze D, Hummel M, Leser U. Comparative assessment of differential network analysis methods. Brief Bioinform. 2016;bbw061.
34. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R; with the HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Med Inform Decis Mak. 2017 Apr 10;17(1):36.
35. Tang LA, Johnson KB, Kumah-Crystal YA. Breadcrumbs: Assessing the feasibility of automating provider documentation using electronic health record activity. AMIA Annu Symp Proc. 2018;2018:1008–17.