Using Time Series Clustering to Segment and Infer Emergency Department Nursing Shifts from Electronic Health Record Log Files

Amanda J. Moy, MPhil, MPH, MA;[1] Kenrick D. Cato, RN, PhD;[2,3] Jennifer Withall, RN, PhD;[3] Eugene Y. Kim, MD;[2] Nicholas Tatonetti, PhD;[1] Sarah C. Rossetti, RN, PhD[1,3]
[1]Columbia University Department of Biomedical Informatics, NY, NY, USA; [2]Columbia University Irving Medical Center Department of Emergency Medicine, NY, NY, USA; [3]Columbia University School of Nursing, NY, NY, USA

Abstract

Few computational approaches exist for abstracting electronic health record (EHR) log files into clinically meaningful phenomena like clinician shifts. Because shifts are a fundamental unit of work recognized in clinical settings, shifts may serve as a primary unit of analysis in the study of documentation burden. We conducted a proof-of-concept study to investigate the feasibility of a novel approach using time series clustering to segment and infer clinician shifts from EHR log files. From 33,535,585 events captured between April-June 2021, we computationally identified 43,911 potential shifts among 2,285 (74.2%) emergency department nurses. On average, computationally-identified shifts were 10.6±3.1 hours long. Based on data distributions, we classified these shifts based on type: day, evening, night; and length: 12-hour, 8-hour, other. We validated our method through manual chart review of computationally-identified 12-hour shifts achieving 92.0% accuracy. Preliminary results suggest unsupervised clustering methods may be a reasonable approach for rapidly identifying clinician shifts.

Introduction

Clinician documentation burden is recognized as a pervasive problem within our health care system and a major driver of clinician burnout.[1,2] Prior to the broad implementation of electronic health records (EHRs), the primary purpose of clinical documentation was to support direct patient care and to communicate clinical decision-making among care team members. With the digitization of clinical data capture through information systems and simultaneous motivation to enforce regulatory requirements and standards for clinical practice, EHR design and development has unintentionally pivoted the focus of clinician documentation to reimbursement and reporting.[3] In addition to the overall increase of documentation volume and information consumption contributing to documentation burden among clinicians, extensive literature has demonstrated that suboptimal EHR system design, usability, and integration of EHRs[4] in clinical workflows *alone* are associated with more time spent on clinical documentation and burnout among clinicians.[5–7] These concerns have become particularly salient during the COVID-19 pandemic, which has been accompanied by an exodus of healthcare workers from clinical medicine.[8,9]

Mitigating EHR documentation burden will rely on consistent and ongoing evaluation of EHR optimizations[10]; however, suitable, universally agreed upon quantitative and computational methods to examine documentation burden and its proxy, clinician work, remain elusive. Concretely, few standardized, cost-effective and scalable approaches exist to measure and understand the impact of EHR design and usability on clinician workflows.[11,12] Historically, research on clinician workflows have employed time-motion studies, but the growing ubiquity of EHRs and subsequent exponential growth in actively (e.g., patient care-related) and passively (e.g., EHR metadata) collected EHR data have shifted the emphasis from direct observation techniques to secondary data analyses,[11] which have promoted enormous opportunities for health outcomes, health care and health services research. EHR log files are one type of passively collected data which emerged through HIPAA and Meaningful Use mandates requiring that healthcare organizations implement procedures to monitor privacy and security (e.g., auditing and access) of protected health information in computerized information systems.[13] While promising for investigating EHR use and clinician workflows in the EHR, EHR log files are exceedingly convoluted and were not intended for use in research,[18] and insufficient criteria have been established to standardize its data granularity and the breadth of which EHR-related behaviors are captured among its users.[14]

Currently, no prescriptive guidelines exist to organize EHR log files into readily identifiable, clinically meaningful phenomena. Rigorous approaches for abstracting EHR log files into usable, reliable, and actionable abstractions such as clinician activities and discrete shifts—a widely applied primary unit of analysis for clinician work—to examine clinician behaviors are wanting. While many studies that utilize EHR log files investigate clinician workload and workflows at the shift level, limited research has been devoted to developing shift detection methods that singularly rely on EHR log files as a primary data source, can be leveraged to further EHR workflow analyses, and are generalizable to a broad range of clinical contexts. Clinician scheduling data typically reside on administrative systems and databases that are autonomous from EHRs and often require additional permissions and approvals to access as a safeguard for employee privacy. Considerable variation in the sophistication and architecture of these scheduling systems across institutions, settings, sites and clinician roles contribute to the overall variability in their data output structures, which frequently do not exist in machine computable formats (e.g., word-processing documents and portable document format files), and oftentimes require additional data normalization.[15–17] Prior EHR log file studies on shift level work among clinicians have been small-scale, manually identify shifts,[18] rely on auxiliary administrative data, employ heuristics[19] and/or examine aggregate data over extensive time periods (e.g., year).[20] To our best knowledge, only one large-scale EHR log file study has been conducted on automated shift detection among pediatric in-patient residents using rule-based classification, a technique that classifies data according to explicit, human knowledge-based conditions[21]; no studies have explored unsupervised methods, which uncover underlying patterns based on inherent data structures.

It has become increasingly apparent that clinician workflows and work patterns (and subsequently, the experience of documentation burden) vary individually as well as across roles, settings, and institutions; thus, more versatile methods to identify shifts and other clinically meaningful constructs in its study may be warranted.[2] This proof-of-concept study is aimed at investigating the feasibility and scalability of a novel two-step approach for automatically segmenting and inferring clinician shifts from EHR log files (hereinafter referred to as "computationally-identified shifts") using density-based spatial clustering of applications with noise (DBSCAN). DBSCAN is one type of time series clustering method that has been used in a variety of healthcare contexts for medical image segmentation,[22-,25] recommender systems[26] and anomaly detection.[27] DBSCAN groups interesting subsequences of dynamic, time-dependent data[28]—operationalized as "EHR log file timestamps" in this study—of varying time periods and cadences into similar clusters which we define as "shifts". Given the paucity of research in more advanced EHR log file analyses, and variable nature of shift types and lengths among clinicians, the overall motivation of this study is to determine whether time series clustering (i.e., data-driven) approaches are worthwhile to explore for the large-scale automated detection of clinician shifts using EHR log file data. Prior studies on shift level EHR work and clinician workflows using EHR log files have primarily focused on physicians in inpatient and ambulatory settings;[15,16,20,21] few have examined the emergency department (ED).[17] To our knowledge, the majority of research examining shift level nursing work have been undertaken in intensive and acute care units,[19,29] whereas none have investigated ED nursing shifts. Contrasting traditional ambulatory and inpatient settings, ED environments (and shifts) are highly dynamic and largely driven by patient census and acuity.[30] In this paper, we describe our automated shift detection approach, share the results of our algorithm's findings, and validate our methodology using manual chart review. And while we focus on automatically detecting shifts among nurses in the ED in this analysis, we envision further extending our methodology to shifts that are more irregular and difficult to segment, such as ED resident and attending shifts in the future.

Methods

We conducted a large-scale retrospective cohort analysis to automatically detect ED nurse shifts using an unsupervised clustering approach to segment EHR log files and infer shifts based on the "similarity" of EHR log file timestamps for our proof-of-concept study. We extracted raw EHR log files between September 1, 2020 to August 31, 2021 generated from one ED site of a large academic medical center in the northeastern United States for all registered nurses who worked in the ED. This ED site is a large, quaternary care center with over 100,000 patient encounters annually. The institution uses the Epic Systems EHR (Verona, WI, Epic Systems), which was implemented in February 2020. EHR log files were extracted from the Epic Clarity database and contained information on the following five data elements: *process id* (i.e., unique cache process per user login instance), *clinician identification number* (i.e., individual clinician interacting with the EHR), *patient identification number* (i.e., medical record accessed), *event name* (i.e., action performed in the EHR *and* captured in the EHR log files, such as "Flowsheets viewed", "Storyboard viewed", and "MAR administration accepted"), and *timestamp* of when an event was performed. Our final analytical dataset comprised of three months of data (April to June 2021). For completeness, all log events associated with both a patient identification number and an encounter number were included in the analysis regardless of whether they were user- or system-generated. All data were deidentified prior to initiating analysis.

Clustering Algorithms and DBSCAN

Clustering is an unsupervised machine learning task employed for uncovering natural groupings among similar data objects.[22] Based on the distribution of a dataset and grouping strategy, various categories of clustering algorithms can be employed: partitioning, hierarchical, density-based, model-based and grid-based.[23] Traditionally, clustering algorithms have been applied on static data (i.e., features with values that remain constant or change marginally over time)[22]; however, exponential growth of time series data (defined as "a series of data points in similar time spaces"[22]) in last decade has led to increased attention to time series clustering. According to Zolhavarieh et al., time series clustering partitions "interesting subsequence[s] of time series data in[to] the same cluster"[22] to explore underlying patterns and structures.[23] Three categories of time series clustering exist: whole time series clustering, subsequence time series clustering, and time point clustering.[22]

DBSCAN is one type of clustering algorithm which is: (a) density-based (i.e., defines clusters as contiguously dense regions in space separated by low-density regions), and (b) relies on time point clustering (i.e., groups time points based on temporal proximity and similarity of their corresponding values).[22] Specifically, DBSCAN efficiently utilizes the density of data points in space (i.e., neighboring points) to form clusters of arbitrary shapes (i.e., different sizes and densities),[23] and effectively processes outliers that lie in low-density regions. Unlike other clustering algorithms, such as k-means, DBSCAN works well with low dimensional data and is a non-parametric algorithm (i.e., does not require an a priori determination of cluster numbers).[28]

The DBSCAN algorithm requires two parameters: (a) epsilon *(*ε), the furthest distance at which a point will select a neighboring point to form a cluster, and (b) minimum points (n_p), the threshold number of neighboring points required to be considered a valid cluster (i.e., dense region) or alternatively, an outlier (i.e., a data point that statistically differs significantly from other points).[33] While adjusting ε is driven by the distribution of the points in a dataset, n_p is determined by expert knowledge of the domain problem.

Data Preprocessing

Given prior evidence that nurse experience impacts their EHR documentation habits,[31] we assume individual nurse EHR behaviors and workflow patterns were unique and performed independent of other nurses; therefore, we implemented the DBSCAN algorithm at the individual-nurse level. As such, among each individual nurse, DBSCAN may identify one or more clusters of "similar" timestamps, with each unique timestamp cluster potentially representing what we conventionally recognize as individual shifts. Partitioning DBSCAN clustering based on this natural separability (i.e., *clinician identification number*) between successive timestamps in raw EHR log files among individuals additionally accounted for DBSCAN implementation requiring that feature values be continuous (Figure 1). We preprocessed raw EHR log file data by converting all event timestamps to integers which represented total seconds after elapsed UNIX epoch time (i.e., January 1, 1970). We standardized these integers by subtracting the mean and scaling the unit variance such that the data followed a normal distribution N(0,1); these data were normalized so that all feature values resided on the same scale.[32] The data were then sorted by ascending *clinician identification number* and *timestamp*.

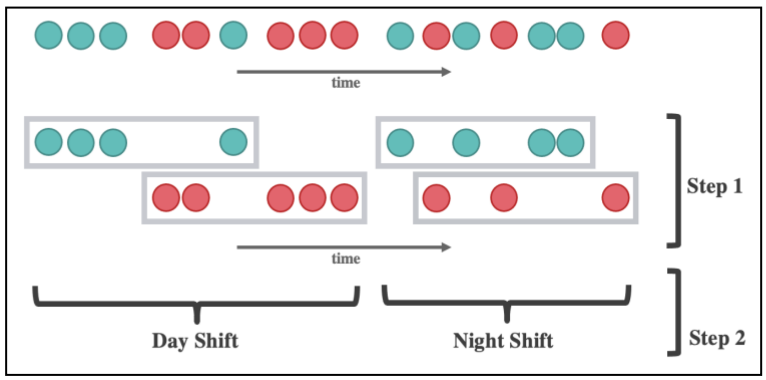

Figure 1. *Schematic of our two-step time series clustering approach to automatically identify shifts using raw EHR log files. In this scenario, EHR event timestamps for two nurses, depicted as spheres (teal and red respectively), are subsetted at the individual-nurse level and clustered into shifts based on timestamp density using DBSCAN (gray boxes) in Step 1. In Step 2, rule-based logic conditioned on the first timestamp identified in the time series is applied to classify clusters into day, evening and night shifts.*

Time Series Clustering of EHR Event Timestamps

We selected one month of data (April 2021) to establish baseline values for parameters ε and n_p, and to benchmark the algorithm's runtime. Initial parameters were determined iteratively using a combination of the following three methods: plots, model diagnostics and manual review.

We explored the optimal ε by assessing elbow plots[25] and executing DBSCAN at the individual-nurse level on a random 1.0% subset of nurses (n=30); elbow plots did not yield practical results with increasing granularity of ε. In addition, we examined individual-nurse level DBSCAN model diagnostics: number of clusters identified, number of outliers (i.e., noise points) and the silhouette index (*SI*). The *SI* is a measure of cluster cohesion and separation (i.e., space between clusters) and represents a mean of all instances *i* of silhouette indices and is calculated using the mean intra-cluster distance (a), and the mean nearest-cluster distance (b) for each sample i or $[b_i - a_i/max(a_i, b)]$.[34] The *SI* ranges between -1 and +1, positive one indicating that within cluster samples are closer to each other and further from other clusters, zero indicating proximity to cluster boundaries, and negative one indicating the potential presence of misclustering;[35] naturally, calculation of *SI* requires that identified cluster counts are ≥2 and less than the total number of samples i in the dataset.[36] Specifically, we manually assessed the quality of clusters by investigating all individual-nurse level model diagnostics for clusters with a *SI* <0.5. We adjusted ε based on the overall *SI* (i.e., mean *SI* among all individual-nurse level *SI*s) as well domain knowledge on nurse workflows from three co-authors (SC, KC, JW) who are also nurse informacists. This tuning process continued until there was a less than 50.0% improvement among *SI* scores that changed. We selected the algorithm with the optimal overall *SI*. Lastly, we generalized the selection of ε to the full 3-month analytical dataset. The optimal model was fixed at ε = 0.015.

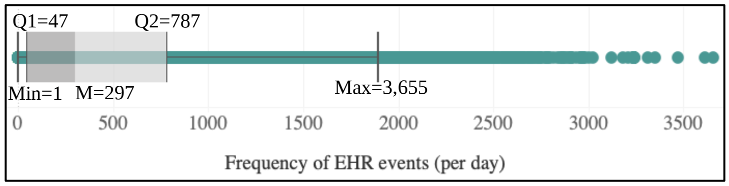

Figure 2. *Distribution of EHR events that nurses logged per day, April 2021 (M=median; Q=quartile; Min=minimum; Max=maximum)*

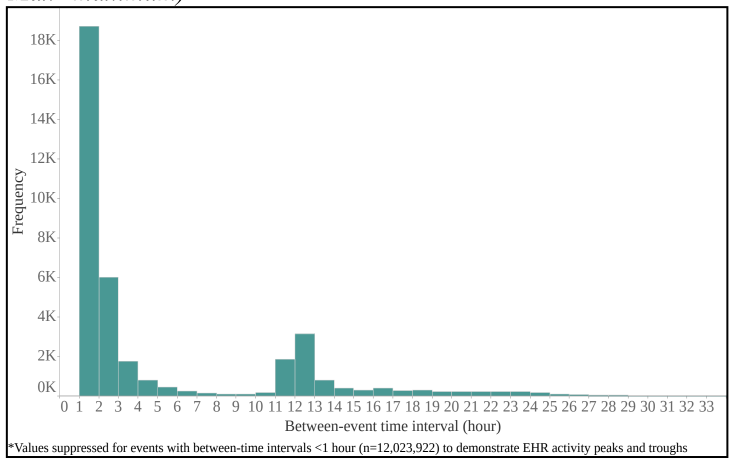

Figure 3. *Distribution of time intervals between EHR events logged among nurses in hours, April 2021**

In this study, n_p is operationalized as the minimum number of EHR events that a nurse must engage in to qualify as a shift. We roughly approximated the number of EHR events nurses engaged in per shift by generating boxplots to examine the distribution of EHR events logged among nurses per day at the individual level (Figure 2). Longitudinally, we examined EHR event trends by plotting volume of EHR events per hour over time for each nurse. Based on the distributions of these plots (see *Q1*), we fixed n_p at 50 EHR events (*Median*=297; $\bar{x}$=500; SD=554.2). Lastly, we plotted a histogram of the between-event intervals (i.e., duration between sequential event timestamps) to understand trends of activity and inactivity among nurses throughout their day. We found that most EHR events were logged within two hours of the previous event, which exponentially decreases until 11 to 13 hours later when EHR events peak again indicating activity in their next shift (Figure 3). As described previously, the final DBSCAN algorithm (ε= 0.015, n_p= 50) was executed on a for-loop at the individual-nurse level, with k-dimensional (k-d) tree optimization.

Rule-based Classification of EHR Event Timestamp Sequence Clusters

After individual nurse EHR event timestamps were computationally segmented into discrete clusters (i.e., shifts), we calculated shift length (i.e., *duration*) by sorting within-cluster EHR event timestamps in ascending order and computing the difference between the first (i.e., shift *start* time) and last timestamps (i.e., shift *end* time) identified in the timestamp sequences for all clusters identified among each nurse (Figure 1). We plotted histograms for shift *duration* and shift *start* time in hours (hrs) respectively for all clusters identified. Based on the distribution of *durations* (Figure 4a), we classified computationally-identified nurse shifts into shift lengths using rule-based logic according the following cut-offs: (a) ≥11hrs and <16hrs (*12-hour* shift); (b) ≥7hrs and <11hrs (*8-hour* shift); and (c) <7hrs or

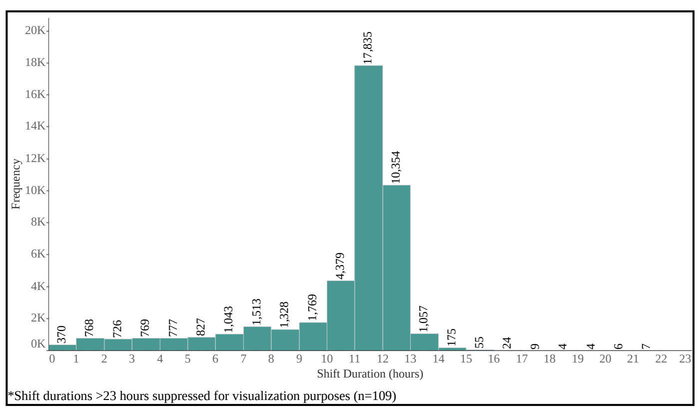

Figure 4a. *Distribution of shift duration among computationally-identified nurse shifts**

≥16hrs (*other*). The majority of computationally-identified shifts were between 10 to 13 hours long (74.2%). Based on the results of shift *start* times (Figure 4b), we also classified *day*, *evening*, and *night* shifts using rule-based logic according to the following start hours: (a) ≥4am and <2pm (*day*); (b) ≥2pm and <6pm (*evening*); (c) ≥6pm and <4am (*night*). Among computationally-identified shifts, most *start* times fell between 7AM and 9AM (44.4%) as well as between 7PM and 9PM (39.3%).

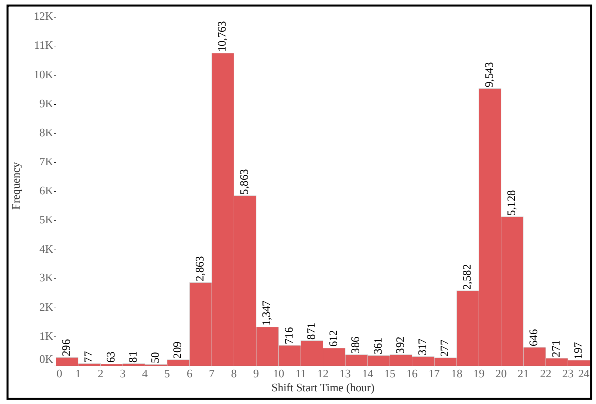

Figure 4b. *Distribution of shift start times among computationally-identified nurse shifts*

Validation of Results

Presently, no gold standard approach exists for validating clinician shifts identified through raw EHR log files; however, prior studies have utilized clinician self-reported shifts[15] and manual chart review.[21] We validated our computationally-identified nurse shifts through a manual chart review of EHR data by one co-author (KC) of a random 10.0% sample of all nurses with at least one computationally-identified, 12-hour shift (n=179). Among each of those nurses, we randomly sampled two computationally-identified, 12-hour shifts for review (n=358); if fewer than two computationally-identified, 12-hour shifts were present in that nurse, we continued sampling nurses until our sample size was achieved. We employed this approach as a large proportion of nurses working at this ED site were travel nurses, float nurses and/or nurse consultants who did not have a primary appointment in the ED, but whose EHR log file data were indistinguishable from permanent ED staff nurses. This sample size has been used in studies of similar context.[21] We assessed the accuracy of our approach by dividing the overall number of computationally-identified, 12-hour shifts that were deemed correct based on chart review by the total number of computationally-identified, 12-hour shifts sampled. All analyses were implemented using Python version 3.8.5 with scikit-learn.[36]

Results

Between April 1, 2021 and June 30, 2022, EHR log files captured 33,535,585 events and their associated timestamps among 3,079 ED nurses who interacted with the EHR. Among those event timestamps, we computationally identified 43,911 potential shifts among 2,285 (74.2%) nurses using DBSCAN. The optimal parameters (ε = 0.015, n_p= 50) yielded the best overall *SI* (0.820). Over 80.0% of nurses had at least one event timestamp classified as an outlier (86.1±125.0 timestamp points; *Median* = 44.0); however, outliers represented a small proportion (0.6%) of the overall data (Table 1).

Table 1. *Aggregate summary of DBSCAN model diagnostics among 3,091 RNs, April-June 2021*

	Model 1 [ε = 0.010, n_p= 50]		**Model 2** [ε = 0.0125, n_p= 50]		**Model 3** [ε = 0.015, n_p= 50]	
Diagnostic	**n (%)**		**n (%)**		**n (%)**	
Number of nurses with ≥1 computationally-identified shifts	2,284 (74.2)		2,284 (74.2)		2,285 (74.2)	
Number of shifts computationally-identified	43,971 (na)		43,983 (na)		43,911 (na)	
Number of event timestamps identified as outliers	225,375 (0.7)		216,165 (0.6)		212,645 (0.6)	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**	**Mean (SD)**	**Range**
Average silhouette index across all nurses	0.817 (0.161)	[-0.443, 0.999]	0.819 (0.159)	[-0.408, 0.999]	0.820 (0.159)	[-0.406, 0.999]

As demonstrated in Figure 4a, the *durations* among computationally-identified shifts were left skewed with large outliers on the right hand side. On average, computationally-identified shifts were 10.6±3.1hrs long (*Median* = 11.6; *IQR* = 1.7; Range[0, 85.6]). After applying rule-based shift classification

Table 2. *Frequency table of shift type by shift length (duration) among computationally-identified shifts.*

	Total Shifts	**EHR Log Events (per Shift)**		
Shift Type	**n (%)**	**Mean (SD)**	**Median**	**Range**
Day	15,089 (51.3)	989.7 (571.3)	929	[50, 4235]
Night	14,337 (48.7)	928.0 (553.8)	848	[50, 3980]
Overall	29,426 (100.0)	959.6 (563.7)	890	[50, 4235]

logic to computationally-identified shifts (see *Methods*), we found that the majority of ED nurse shifts were 12 hours in length (67.1%) followed by 8 hours (20.5%) (Table 2). While 12-hour shifts were largely uniformly distributed between day (51.2%) and night (48.6%) shifts, nearly two-thirds of 8-hour shifts were day shifts. On average, more EHR events were captured per shift among day (989.7±571.3) compared to night (928.0±553.8) shift nurses (Table 3).

Validation Results

Based on our sampling strategy, we ultimately extracted 276 computationally-identified, 12-hour shifts from 179 ED nurses for manual review. Among the 276 computationally-identified day and night shifts, 92.0% were correctly classified as 12-hour day and night shifts. Twenty computationally-identified shifts appeared to have shift *start* and *end* times that were shifted by 12 hours (i.e., misclassification of night and day shifts based on events in EHR log files). The status of two computationally-identified shifts from one nurse could not be verified as not enough information was available among individual patient charts to determine clinician EHR activity; however, based on the EHR log files, over 1,000 events were logged for each of those computationally-identified shifts.

Table 3. *Distribution of the frequency of EHR events that nurses logged per shift, day and night shift only*

Shift Type	Shift Length			
	12-hour	8-hour	Other	Total
	n (%)	n (%)	n (%)	n (%)
Day	15,089 (63.7)	5,720 (24.2)	2,871 (12.1)	23,680 (100.0)
Evening	50 (3.7)	125 (9.3)	1,172 (87.0)	1,347 (100.0)
Night	14,337 (75.9)	3,144 (16.4)	1,403 (7.4)	18,884 (100.0)
Overall	29,476 (67.1)	8,989 (20.5)	5,446 (12.4)	43,911 (100.0)

Discussion

To our best knowledge, this is the second study investigating automated approaches for detecting clinician shifts using EHR log files,[21] and the first to implement an unsupervised time series clustering method. In this study, we computationally identified 43,911 potential shifts among 2,285 (74.2%) ED nurses using DBSCAN. On average, computationally-identified shifts were 10.6±3.1 hours long. Validation of computationally-identified, 12-hour day and night shifts *(*n=276) among ED nurses through retrospective chart review demonstrated that our two-step approach detected 12-hour shifts at a 92.0% accuracy; however, this may be higher as retrospective chart reviews are complex and prone to error (~10.0%).[37] For example, among two computationally-identified, 12-hour shifts that could not be ascertained and validated in the EHR charts, both consistently logged event timestamps (n >1,000) over the course of 12 hours, suggesting high volume clinician EHR activity. In contrast, Dziorny et al.[21] found no mismatches between scheduled shifts and "EHR-calculated" shifts using their rule-based classification method. It is worthwhile to note that Dziorny and colleagues[21] obtained EHR log files using trainee names, whereas we inclusively examined all available data for the specified time period. Further investigation is needed to discern the source of the incongruity between EHR log files and EHR charts in our study (e.g., data access issues or system bugs).

Overall, our findings are consistent with our knowledge of ED nursing shift structures at the study site, which are typically 12 hours long and implements staggered start times centered at 7AM and 7PM, respectively (see Figures 4a & 4b). This suggests *data-driven methods* may be a reasonable approach to examine shifts—and more generally, clinician workflows—in EHR log files as an alternative to traditional rule-based algorithms (RBAs) for some research questions. For example, although RBAs are human interpretable, manual generation of rules are often difficult and time-consuming to perform, and may lack generalizability and scalability to other contexts (e.g., distinct clinician roles and types, practice settings and institutions) including evolving contexts (e.g., changes in the EHR design or business process requirements). Frequently, RBAs are specialized for a specific problem and require extensive domain knowledge. This lack of portability among RBAs is readily apparent in their fixed thresholds, rendering them incapable of independently learning "new" rules, which must be incrementally added or removed, and requiring continuous data (e.g., time) to be discretized. As a result, RBAs may not be appropriate for some tasks involving highly dynamic and continuously generated data, such as EHR log files. For instance, among the 254 computationally-identified, 12-shifts that we correctly validated, nearly 40.0% were longer than 12 hours and none were exactly 12 hours long. Strict RBAs may artificially truncate these actual durations spent in the EHR within one scheduled shift depending on the rules established for the algorithm, and therefore, underrepresent and underestimate actual clinician interactions and work in the EHR (i.e., natural clinician behaviors), and ultimately, level of documentation burden. While typical shift types are day, evening and night, and shift lengths are 12 and 8 hours long, individual clinician shifts may vary across days and lengths—not accounting for overtime shifts. Given this mutability, machine learning algorithms may be a suitable alternative for detecting clinician shifts as they optimally solve “general questions,” are simple to use and implement, and update automatically when presented with novel data.

In the context of automated shift detection, DBSCAN specifically, may be a superior approach compared to alternative clustering algorithms (e.g., k-means) as it does not require a priori knowledge of cluster numbers and is capable of efficiently grouping data points with arbitrary shapes and sizes.[23] This flexibility represents a key advantage in the domain of shift detection as: (a) actual clinician shifts are indiscernible based on raw EHR log files alone, and (b) actual clinician shifts may not abide by customary scheduled shift lengths (e.g., 12-hour, 8-hour, etc.) which (in reality) may be longer or shorter depending on whether EHR documentation was completed before the end of a scheduled nursing shift, or after the change of shift handover. Additionally, because density-based algorithms like DBSCAN partition high-density data regions from low-density data regions, they can detect (and flag) outliers within a task, such as sporadic interactions in the EHR not associated with work during a scheduled shift.[3] Longitudinal examination of individual EHR event trends among nurses over time (i.e., volume of EHR events per hour) in our study demonstrates that some nurses document voluminously over short time periods (e.g., two hours) and then abruptly stop for several hours; these documentation habits may represent nurses that hold managerial positions in the ED. Based on the results of our algorithm, a marginal proportion of outliers (0.6%) were identified in our dataset, suggesting that ED nurses rarely engage in work outside their scheduled shift. While consistent with the results of our ongoing qualitative study on ED nurse documentation burden in the EHR,[41] further investigation is required to ensure that these outlier EHR event timestamps were not misclassified and that algorithm parameters were not misspecified as adjusting parameters ε and n_p would determine whether those EHR event timestamps are included or excluded in the analysis; as such, there are tradeoffs to DBSCAN. For example, too stringent parameters may lead to valid EHR interactions within a shift being divided into distinct shifts or recognized as outliers (and therefore missed), whereas too lax parameters may lead to the agglomeration of EHR interactions among multiple independent shifts into a single broad shift.

Inferring clinician shifts from raw EHR log files remains a difficult undertaking. In this study we classified computationally-identified shifts into shifts and types using expert knowledge-driven heuristics for descriptive purposes; however, our intention is to holistically and longitudinally examine computationally-identified shifts without these ascribed labels with the goal of understanding factors that drive unusually longer than expected time in the EHR (e.g., nature and type of work conducted). As described previously, EHR interactions outside a scheduled shift may have led to the misclassification of these ascribed shift types and lengths, and may have resulted in the detection of extraneous shifts or the agglomeration of several shifts. Additionally, it is possible that valid EHR event timestamps within a shift may have been misidentified as outliers. These scenarios are conceivable as over 12.0% of computationally-identified shifts were classified as *other* (i.e., <7hrs or ≥16hrs long). These are known challenges that are intrinsic to the use of EHR log files for understanding clinical activity.[11,14] It is likely that location information would improve our algorithm’s overall performance by improving the resolution of high-density areas (i.e., shifts) and reducing the presence of noise (i.e., interaction with the EHR outside the ED) so we can differentiate onsite EHR work performed outside normally scheduled shift time with worked performed at home. Ideally, these location data would be integrated with existing EHR systems and databases; however, interoperability among distinct vendor proprietary systems and tools is slow to evolve, and presently, an unrealistic endeavor. For instance, administrative data on nurse and physician scheduled shifts at this ED site reside on two independent web-based scheduling, timekeeping and attendance systems, *Kronos Workforce Timekeeper* (Kronos) (New York, NY, Ultimate Kronos Group) and *Shift Admin* (Atlanta, GA, Qgenda) respectively, which are not jointly integrated nor integrated with the EHR. Meanwhile, resident scheduling data reside on Google sheets and portable document files, which are not machine computable. These scheduling data are not always up-to-date and/or accurate; and while administrative systems data may provide information on shift scheduling, they may be difficult to access due to privacy concerns. For these reasons, we advocate for the development of standardized, quantitative and computational methods (to investigate clinically meaningful constructs associated with clinical workflows) that do not require data that are not integrated in the EHR.

Future Directions

Validation of our results using shift scheduling data is still in process and will be complete by the end of 2022; these results will be included in future revisions and presentations. Most recently, we randomly sampled 10.0% of all nurses with at least one computationally-identified shift (independent of shift length) detected in the raw data extract (n=230) for validation against data captured in Kronos. Our comprehensive two-step process is summarized in Figure 5.

Because shift length has been used as an objective organizational metric to examine clinician work in the past, future efforts will focus on extending our two-step approach to more irregular shifts and schedules (e.g., overtime)—such as resident and attending shifts—in addition to larger data volumes, which may be more challenging to segment and infer using traditional RBAs. In the context of documentation burden and burnout,[38] these shift structures are critical to examine as irregular shift work are associated with reduced performance[39] and increased risk of error.[40] Moreover, we are beginning to explore whether computationally-identified shifts can be grouped into similar shift types and lengths as well as EHR engagement patterns in an unsupervised fashion (i.e., without the use of rule-based logic); the premise of this exploratory analysis is to help lay the groundwork for methods in EHR workflow analyses, as well as in the phenotyping of clinicians based on EHR interaction habits so that targets to reduce documentation burden can be identified.

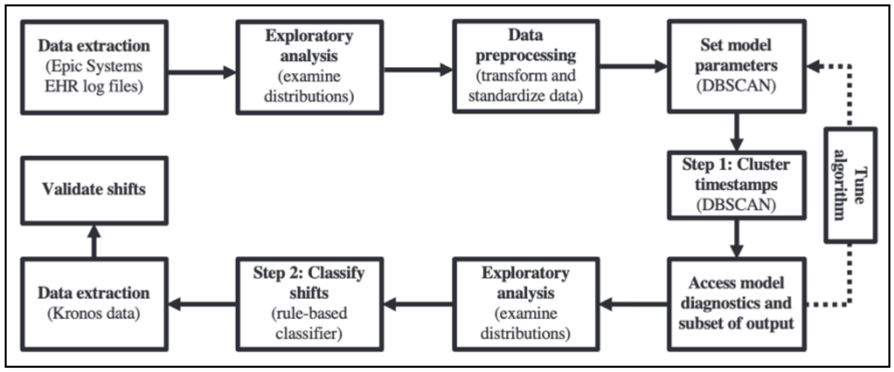

Figure 5. *Process model of our two-step approach starting with raw EHR log file data extraction to the validation of computationally-identified shifts using Kronos data.*

Limitations

This study has several limitations. Because existing EHR log files do not provide information on the location at which nurses interacted with the EHR (e.g., workstation or IP address), we are unable to distinguish whether EHR interactions captured outside of conventional scheduled shift hours occurred contemporaneous to their scheduled shift (i.e., after they clocked out but in the ED) or outside their shift (i.e., in their residence); however, our study team's ongoing qualitative research on EHR documentation burden among ED nurses reveal that few nurses access the EHR and/or document outside the ED setting.[41] While broad variability may exist in EHR interaction habits within and across nurses, [31] we set algorithm parameters (ε, n_p) identically across all individual-nurse level data; those fixed parameter values may not have been generalizable to all nurses. In addition, validation of our method was conducted by only one study team member and for computationally-identified shifts of 12 hours only. These results may be biased towards nurses who engage frequently and predictably within the EHR during their shifts, and therefore, may be nonrepresentative of all nurses and/or shifts. Lastly, we examined three months of data which may be nonrepresentative of all potential shifts that exist.

Conclusion

Developing standardized, novel approaches to abstract clinically meaningful constructs and actionable information from EHR log files will be foundational to impactful analysis on clinician workflows in the EHR, and ultimately, documentation burden. In this study, we implemented and validated a two-step approach that applies time series clustering to automatically segment and infer ED nurse shifts from EHR log files, which yielded a 92.0% accuracy among computationally-identified, 12-hour shifts. Our work is significant as it demonstrates that data-driven methods (as opposed to expert knowledge driven methods) may be utilized to extract meaning (i.e., shift level EHR work) from EHR log files.

Acknowledgements

This research was supported by the U.S. National Library of Medicine training fellowship award 5T15LM007079 and the Agency for Healthcare Research and Quality grant R01HS028454.

References

1. Sittig DF, Wright A, Ash J, Singh H. New Unintended Adverse Consequences of Electronic Health Records. Yearb Med Inform. 2016 Nov 10;(1):7-12. doi: 10.15265/IY-2016-023.
2. Moy AJ, Schwartz JM, Withall J, Lucas E, Cato KD, Rosenbloom ST, et al. Clinician and Health Care Leaders' Experiences with-and Perceptions of-COVID-19 Documentation Reduction Policies and Practices. Appl Clin informaticsfile///Users/amanda/Downloads/25140332.nbib. 2021 Oct;12(5):1061–73.

3. Stehman CR, Testo Z, Gershaw RS, Kellogg AR. Burnout, Drop Out, Suicide: Physician Loss in Emergency Medicine, Part I. West J Emerg Med. 2019 May;20(3):485-494. doi: 10.5811/westjem.2019.4.40970. Epub 2019 Apr 23. Erratum in: West J Emerg Med. 2019 Aug 21;20(5):840-841.
4. Park SY, Lee SY, Chen Y. The effects of EMR deployment on doctors' work practices: a qualitative study in the emergency department of a teaching hospital. Int J Med Inform. 2012 Mar;81(3):204-17. doi: 10.1016/j.ijmedinf.2011.12.001. Epub 2012 Jan 3.
5. Melnick ER, Dyrbye LN, Sinsky CA, Trockel M, West CP, Nedelec L, et al. The Association Between Perceived Electronic Health Record Usability and Professional Burnout Among US Physicians. Mayo Clin Proc. 2020 Mar;95(3):476-487. doi: 10.1016/j.mayocp.2019.09.024. Epub 2019 Nov 14. PMID: 31735343.
6. Tajirian T, Stergiopoulos V, Strudwick G, Sequeira L, Sanches M, Kemp J, et al. The Influence of Electronic Health Record Use on Physician Burnout: Cross-Sectional Survey. J Med Internet Res. 2020 Jul;22(7):e19274.
7. Melnick ER, West CP, Nath B, Cipriano PF, Peterson C, Satele D V, et al. The association between perceived electronic health record usability and professional burnout among US nurses. J Am Med Informatics Assoc [Internet]. 2021;28(8):1632–41. Available from: https://doi.org/10.1093/jamia/ocab059
8. Holmgren AJ, Downing NL, Tang M, Sharp C, Longhurst C, Huckman RS. Assessing the impact of the COVID-19 pandemic on clinician ambulatory electronic health record use. J Am Med Informatics Assoc [Internet]. 2021;29(3):453–60. Available from: https://doi.org/10.1093/jamia/ocab268
9. Prasad K, McLoughlin C, Stillman M, Poplau S, Goelz E, Taylor S, et al. Prevalence and correlates of stress and burnout among U.S. healthcare workers during the COVID-19 pandemic: A national cross-sectional survey study. EClinicalMedicine. 2021 May;35:100879.
10. Shah T, Kitts AB, Gold JA, Horvath K, Ommaya A, Frank O, et al. Electronic Health Record Optimization and Clinician Well-Being: A Potential Roadmap Toward Action. NAM Perspect. 2020 Aug 3;2020:10.31478/202008a. doi: 10.31478/202008a.
11. Moy AJ, Schwartz JM, Chen R, Sadri S, Lucas E, Cato KD, et al. Measurement of clinical documentation burden among physicians and nurses using electronic health records: a scoping review. J Am Med Inform Assoc. 2021 Apr 23;28(5):998-1008. doi: 10.1093/jamia/ocaa325.
12. Ratwani RM, Savage E, Will A, Arnold R, Khairat S, Miller K, et al. A usability and safety analysis of electronic health records: a multi-center study. J Am Med Inform Assoc. 2018 Sep;25(9):1197–201.
13. 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 Jun 2;172(11 Suppl):S116-S122. doi: 10.7326/M19-0871.
14. Rule A, Chiang MF, Hribar MR. Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods. J Am Med Inform Assoc. 2020 Mar 1;27(3):480-490. doi: 10.1093/jamia/ocz196.
15. Cox ML, Farjat AE, Risoli TJ, Peskoe S, Goldstein BA, Turner DA, Migaly J. Documenting or Operating: Where Is Time Spent in General Surgery Residency? J Surg Educ. 2018 Nov;75(6):e97-e106. doi: 10.1016/j.jsurg.2018.10.010.
16. Long CP, Tai-Seale M, El-Kareh R, Lee JE, Baxter SL. Electronic Health Record Use among Ophthalmology Residents while on Call. J Acad Ophthalmol. 2020 Jul;12(2):e143–50.
17. Kannampallil TG, Denton CA, Shapiro JS, Patel VL. Efficiency of Emergency Physicians: Insights from an Observational Study using EHR Log Files. Appl Clin Inform. 2018 Jan;9(1):99-104. doi: 10.1055/s-0037-1621705. Epub 2018 Feb 7.
18. Shine D, Pearlman E, Watkins B. Measuring Resident Hours by Tracking Interactions with the Computerized Record. Am J Med [Internet]. 2010;123(3):286–90. Available from: https://www.sciencedirect.com/science/article/pii/S000293430901064X
19. Womack DM, Hribar MR, Steege LM, Vuckovic NH, Eldredge DH, Gorman PN. Registered Nurse Strain Detection Using Ambient Data: An Exploratory Study of Underutilized Operational Data Streams in the Hospital Workplace. Appl Clin Inform. 2020 Aug;11(4):598–605.
20. Wang JK, Ouyang D, Hom J, Chi J, Chen JH. Characterizing electronic health record usage patterns of inpatient medicine residents using event log data. PLoS One. 2019 Feb 6;14(2):e0205379. doi: 10.1371/journal.pone.0205379.
21. Dziorny AC, Orenstein EW, Lindell RB, Hames NA, Washington N, Desai B. Automatic Detection of Front-Line Clinician Hospital Shifts: A Novel Use of Electronic Health Record Timestamp Data. Appl Clin Inform. 2019 Jan;10(1):28–37.
22. Zolhavarieh S, Aghabozorgi S, Teh YW. A review of subsequence time series clustering. Scientific World

Journal. 2014;2014:312521.
23. Javed A, Lee BS, Rizzo DM. A benchmark study on time series clustering. Mach Learn with Appl [Internet]. 2020;1:100001. Available from: https://www.sciencedirect.com/science/article/pii/S2666827020300013
24. Baselice F, Coppolino L, D'Antonio S, Ferraioli G, Sgaglione L. A DBSCAN based approach for jointly segment and classify brain MR images. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015:2993–6.
25. Silva WS, Jasbick DL, Wilson RE, Azevedo-Marques PM, Traina AJM, Santos LFD, et al. A Two-Phase Learning Approach for the Segmentation of Dermatological Wounds. In: 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS). 2019:343–8.
26. Sajde M, Malek H, Mohsenzadeh M. RecoMed: A Knowledge-Aware Recommender System for Hypertension Medications. 2022;30:1-9.
27. Lin CH, Hsu KC, Johnson KR, Luby M, Fann YC. Applying density-based outlier identifications using multiple datasets for validation of stroke clinical outcomes. Int J Med Inform. 2019 Dec;132:103988. doi: 10.1016/j.ijmedinf.2019.103988. Epub 2019 Oct 3.
28. Vijayalaksmi S, Punithavalli M. A Fast Approach to Clustering Datasets using DBSCAN and Pruning Algorithms. Int J Comput Appl. 2012;60:1–7.
29. Collins S, Couture B, Kang MJ, Dykes P, Schnock K, Knaplund C, et al. Quantifying and Visualizing Nursing Flowsheet Documentation Burden in Acute and Critical Care. AMIA. Annu Symp proceedings AMIA Symp. 2018;2018:348–57.
30. Moy AJ, Aaron L, Cato KD, Schwartz JM, Elias J, Trepp R, et al. Characterizing Multitasking and Workflow Fragmentation in Electronic Health Records among Emergency Department Clinicians: Using Time-Motion Data to Understand Documentation Burden. Appl Clin Inform. 2021 Oct;12(5):1002–13.
31. Guo W, Kim JH, Smith B, Despins L. How Nurse Experience Influences the Patterns of Electronic Medical Record Documentation in an Intensive Care Unit. Proc Hum Factors Ergon Soc Annu Meet [Internet]. 2019;63(1):708–12. Available from: https://doi.org/10.1177/1071181319631052
32. Liço L, Enesi I, Çiço B. Analyzing Performance of Clustering Algorithms on a Real Retail Dataset. In: 2021 International Conference on Information Technologies (InfoTech). 2021:1–6.
33. Wang C, Ji M, Wang J, Wen W, Li T, Sun Y. An Improved DBSCAN Method for LiDAR Data Segmentation with Automatic Eps Estimation. Sensors [Internet]. 2019;19(1). Available from: https://www.mdpi.com/1424-8220/19/1/172
34. Li M, Bi X, Wang L, Han X. A method of two-stage clustering learning based on improved DBSCAN and density peak algorithm. Comput Commun [Internet]. 2021;167:75–84. Available from: https://www.sciencedirect.com/science/article/pii/S0140366420320259
35. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math [Internet]. 1987;20:53–65. Available from: https://www.sciencedirect.com/science/article/pii/0377042787901257
36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res [Internet]. 2011;12:2825–30. Available from: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html
37. Feng JE, Anoushiravani AA, Tesoriero PJ, Ani L, Meftah M, Schwarzkopf R, Leucht P. Transcription Error Rates in Retrospective Chart Reviews. Orthopedics. 2020 Sep 1;43(5):e404-e408. doi: 10.3928/01477447-20200619-10. Epub 2020 Jul 7.
38. Caruso CC. Negative impacts of shiftwork and long work hours. Rehabil Nurs. 2014 Jan-Feb;39(1):16-25. doi: 10.1002/rnj.107. Epub 2013 Jun 18.
39. Ganesan S, Magee M, Stone JE, Mulhall MD, Collins A, et al. The Impact of Shift Work on Sleep, Alertness and Performance in Healthcare Workers. Sci Rep. 2019 Mar 15;9(1):4635. doi: 10.1038/s41598-019-40914-x.
40. Clendon J, Gibbons V. 12 h shifts and rates of error among nurses: A systematic review. Int J Nurs Stud. 2015 Jul;52(7):1231-42. doi: 10.1016/j.ijnurstu.2015.03.011. Epub 2015 Apr 4.
41. Moy AJ, Rossetti SC. Evaluation of clinician clinician documentation burden using EHR log files. 2022. [ongoing project].