

Automatic Detection of Front-Line Clinician Hospital Shifts: A Novel Use of Electronic Health Record Timestamp Data

Adam C. Dziorny¹ Evan W. Orenstein² Robert B. Lindell¹ Nicole A. Hames² Nicole Washington³
Bimal Desai⁴

¹ Division of Critical Care Medicine, Department of Anesthesia and Critical Care, The Children's Hospital of Philadelphia, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States

² Division of Hospital Medicine, Department of Pediatrics, Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, Georgia, United States

³ Pediatrics Residency Program, Department of Pediatrics, The Children's Hospital of Philadelphia, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States

⁴ Division of General Pediatrics, Department of Pediatrics, The Children's Hospital of Philadelphia, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States

Address for correspondence Adam C. Dziorny, MD, PhD, Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia, Roberts 15-193, 2716 South Street, Philadelphia, PA 19146, United States (e-mail: dziornya@email.chop.edu).

Appl Clin Inform 2019;10:28–37.

Abstract

Objective Excess physician work hours contribute to burnout and medical errors. Self-report of work hours is burdensome and often inaccurate. We aimed to validate a method that automatically determines provider shift duration based on electronic health record (EHR) timestamps across multiple inpatient settings within a single institution.

Methods We developed an algorithm to calculate shift start and end times for inpatient providers based on EHR timestamps. We validated the algorithm based on overlap between calculated shifts and scheduled shifts. We then demonstrated a use case by calculating shifts for pediatric residents on inpatient rotations from July 1, 2015 through June 30, 2016, comparing hours worked and number of shifts by rotation and role.

Results We collected 6.3×10^7 EHR timestamps for 144 residents on 771 inpatient rotations, yielding 14,678 EHR-calculated shifts. Validation on a subset of shifts demonstrated 100% shift match and $87.9 \pm 0.3\%$ overlap (mean \pm standard error [SE]) with scheduled shifts. Senior residents functioning as front-line clinicians worked more hours per 4-week block (mean \pm SE: 273.5 ± 1.7) than senior residents in supervisory roles (253 ± 2.3) and junior residents (241 ± 2.5). Junior residents worked more shifts per block (21 ± 0.1) than senior residents (18 ± 0.1).

Conclusion Automatic calculation of inpatient provider work hours is feasible using EHR timestamps. An algorithm to assess provider work hours demonstrated criterion validity via comparison with scheduled shifts. Differences between junior and senior residents in calculated mean hours worked and number of shifts per 4-week block were also consistent with differences in scheduled shifts and duty-hour restrictions.

Keywords

- ▶ electronic health records
- ▶ physicians
- ▶ burnout
- ▶ algorithms

received
August 6, 2018
accepted after revision
November 19, 2018

© 2019 Georg Thieme Verlag KG
Stuttgart · New York

DOI <https://doi.org/10.1055/s-0038-1676819>.
ISSN 1869-0327.

Background and Significance

Excess work hours among inpatient physicians increase the risk of burnout and medical errors. Burnout, or the exhaustion and lack of satisfaction in work, is a substantial problem among all physicians^{1,2} including front-line clinicians in inpatient settings.³ Implementation of electronic health records (EHRs) has changed the workflow of inpatient providers⁴ and has contributed to increased burnout in both the inpatient and outpatient settings.^{5–8} Physicians spending more time on “nonclinical clerical tasks” are at greater risk of burnout.⁸ Recent estimates of total time that trainee physicians spend in direct patient care are as low as 12% of total work hours,⁹ with nonclinical clerical work making up the largest proportion of trainee time.^{10–12} In addition to contributing to burnout,² long shift durations with frequent overnight shifts such as those commonly worked by trainee physicians can contribute to fatigue among providers which may lead to worse clinical performance¹³ and increased medical errors.¹⁴ Accurate measurement of shift duration and EHR use is critical to understanding this burden on providers.

Unfortunately, “gold standard” measurement of shifts using continuous observation time–motion studies is time consuming and costly, and thus not practical in daily operations.^{15–17} We define a shift as the time period during which the clinician is providing patient care demarcated by the time starting work and finishing work. Work-hour measurement among inpatient physicians is especially challenging as it would require on-call or overnight observation.¹¹ Limited time–motion observations were conducted in the iCOMPARE study, but even in this large well-funded trial, observations only began on weekdays and just 4.2% of observations in eligible programs lasted more than 24 hours.⁹ Ideally inpatient work-hour measurement tools would be objective, not subject to recall bias, noninvasive, and relatively inexpensive.

Front-line clinicians spend a substantial amount of time interacting with the EHR.^{9,11,12,18} Most EHR systems record when users perform specific actions for auditing purposes, and secondary use of these timestamp data has been used for clinical workflow analysis in the outpatient setting.^{19–22} Once validated, EHR timestamp data have proven useful for examining specific aspects of the provider workflow, including the impact of trainees in the office,²³ time spent charting after-hours by providers,²⁴ use of care team identification tools,²⁵ and between-clinic-visit outpatient tasks.²⁶ One prior study has demonstrated manual extraction of EHR timestamps to capture resident work hours during a single rotation in the medical intensive care unit.²⁷ However, converting timestamps into shifts to capture work hours requires an understanding of workflow and properties of the EHR timestamp data. A critical need exists in the development of a robust, validated algorithm for converting EHR timestamp data into shifts for inpatient providers.

Objective

In this study we describe a method of calculating inpatient provider shifts from EHR timestamps across multiple inpa-

tient work settings and validate this methodology by comparing scheduled shifts to calculated shifts in a large, quaternary care pediatric hospital. We further demonstrate the potential applicability of this process by extracting pediatric resident trainee inpatient shifts and showing differences across rotations and roles.

Methods

This study was approved by the Institutional Review Board at the Children’s Hospital of Philadelphia.

Collection of EHR Timestamps

The hospital EHR (Epic Systems, Verona, Wisconsin, United States) stores timestamps of events as a record of user interactions with the system. While the level of logging is implementation-dependent, similar logging is available in other installations and from other EHR vendors. At our institution, event timestamps along with the identifier (ID) of the user who triggered each event are written to an IBM enterprise Hadoop Distributed File System (HDFS) running DB2 Big SQL managed by the Information Systems (IS) department. This platform was chosen to archive the EHR timestamps as it provides scalable storage for large structured data with parallel query execution to speed access. To develop and validate our algorithm, a list of all trainee physicians on inpatient units was obtained from the residency program, and their corresponding user names were obtained from the EHR clinical data warehouse. We collected all events logged for trainee physicians who were pediatric residents during the 2015 to 2016 academic year at Children’s Hospital of Philadelphia, a large urban academic freestanding U.S. children’s hospital. Examples of events logged include “system login,” “open patient record,” “access chart review activity,” and “sign order.” Information for each event includes timestamp, user name, workstation, patient, and encounter (if applicable) (see [►Supplementary Table S1](#) [available in the online version] for an example event log). We excluded event timestamps from workstations not in the hospital (e.g., mobile or at-home access), as we focused specifically on inpatient provider shifts.

We developed a custom algorithm to convert an array of discrete timestamps into EHR-calculated shifts ([►Fig. 1](#)) in an iterative process with our training dataset. Importantly we did not filter or limit on specific event types (metric IDs in our EHR) such as “login” and “logout”; all EHR-generated log events were included so as to increase generalizability. First, EHR event intervals were calculated as the difference between each subsequent EHR event timestamp (in seconds) for a given user. Next, shifts were defined by setting the starting timestamp after a “long” EHR event interval (e.g., the first login in the morning, after a night of no EHR activity) and the end timestamp prior to the next “long” EHR event interval (e.g., the last EHR event before another night of no EHR activity). We defined these “long”-duration heuristics as the cut points in our EHR event intervals through a manual review process, with values provided below.

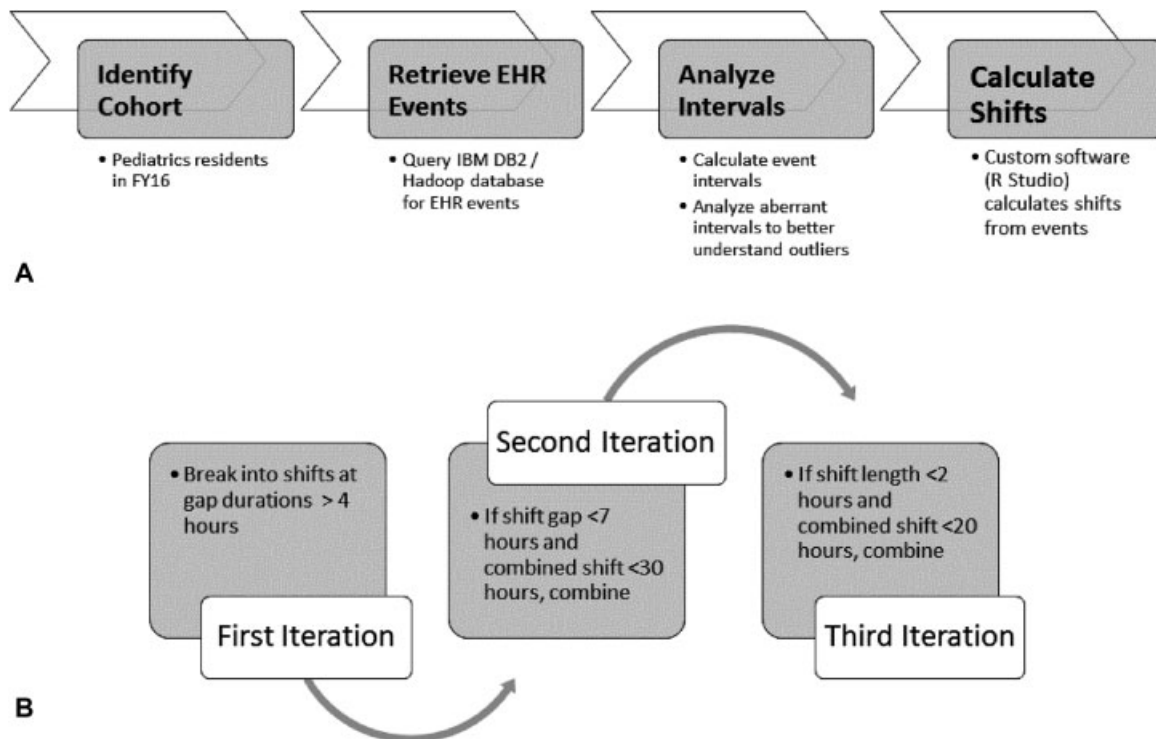


Fig. 1 (A,B) Timestamps were retrieved from the hospital EHR (Epic Systems, Verona, Wisconsin, United States) and processed for shift calculation. Shifts were calculated from discrete event timestamps using a three-pass algorithm which was developed based on knowledge of our clinical workflows. EHR, electronic health record.

Next, to accurately capture all types of shifts (e.g., daytime and overnight), we refined our algorithm in multiple iterations. We targeted our refinement to our validation dataset (see section “Validation to Scheduled Shifts”), with the goal of matching 100% of EHR-calculated shifts to scheduled shifts in the validation set. Cut points were identified by generating histograms of event intervals, exploring this space, and identifying the cut points that resulted in the best shift matching. To inform choices for these heuristic cut points, we also identified workflow elements related to EHR use for pediatric residents through discussion between three current chief residents at the time of algorithm development (A. C.D., R.B.L., and N.A.H.), a current trainee (E.W.O.), and two former chief residents (N.W. and B.D.). Specific elements of trainee workflow that were found to be important included EHR interaction during preroounding and rounding, educational conferences, trainee call schedule, expected time away from the EHR overnight, and expected start and end times for inpatient shifts. After each iteration we reviewed EHR-calculated shifts that did not match scheduled shifts to identify the failure points.

In the first iteration all event intervals with durations >4 hours were broken into separate shifts. This was intentionally chosen to be a short duration, much less than the required time-off between shifts of 10 hours, knowing we might “split” true shifts. Indeed we observed that overnight shifts were sometimes split when no EHR activity was recorded for >4 hours, for example, when the trainee was able to sleep overnight. Therefore, a subsequent iteration combined sequential shifts if the EHR interval was <7 hours

and the combined shift length was <30 hours. Again these were chosen based on workflow knowledge that the minimum time between shifts should not be less than 7 hours, and that overnight shifts with long-duration breaks from the EHR will likely not exceed 30 hours. This accurately combined overnight shifts; however, in rare instances trainees completing work before rounds and then stepping away from the EHR for more than 4 hours were left with short “dangling” shifts in the mornings. Therefore our final iteration combined sequential shifts if one of the shifts had a length <2 hours, as this was unlikely to be an actual scheduled shift, and if the combined shift duration was <20 hours. The resulting final algorithm with these defined heuristic cut points was used for all subsequent analysis.

Validation to Scheduled Shifts

We considered validating trainee-physician shifts against self-reported shifts; however, multiple studies have demonstrated that self-reported hours are not accurate^{28–30} and an informal review of our reporting system suggested fewer than 70% of trainees had self-reported shifts. Therefore, we based our validation on manual review of scheduled shifts.

A subset of trainee physicians was randomly selected for manual review to determine shift-match precision and percent overlap of shift durations. A total of ~10% of shifts were sampled for this validation. Scheduled shifts were recorded from the scheduling website used by the residency program, Amion.com (Newtown, Massachusetts, United States). Shift start and end times were established *a priori* based on shift type and location (e.g., a senior resident on an oncology day

shift was scheduled to start at 7 a.m. and to end at 4 p.m.). EHR-calculated shifts were matched to scheduled shifts manually, and the number of matching and mismatching shifts (either an EHR-calculated shift with no scheduled shift or a scheduled shift with no matching EHR-calculated shift) was reported. Additionally, for all matched shifts we calculated the percent overlap as the time period in common (“overlapping duration”) divided by the time period defined by the earliest start timestamp and the latest end timestamp of either the scheduled or calculated shift (“maximum duration”), as follows:

$$\frac{\min(\text{End}_{\text{Sched}}, \text{End}_{\text{Calc}}) - \max(\text{Start}_{\text{Sched}}, \text{Start}_{\text{Calc}})}{\max(\text{End}_{\text{Sched}}, \text{End}_{\text{Calc}}) - \min(\text{Start}_{\text{Sched}}, \text{Start}_{\text{Calc}})} \times 100$$

In this equation, “Start” and “End” refer to the start and end timestamps of the scheduled (“Sched”) and EHR-calculated (“Calc”) shifts (→ Fig. 2). Lastly, we calculated the number of minutes that EHR-calculated shifts started later than scheduled, and the number of minutes that EHR-calculated shifts ended earlier than scheduled.

Use Case: Pediatric Resident Shifts by Rotation and Role

After manual validation, we applied our algorithm across all pediatric and combined medicine-pediatric residents practicing on pediatric inpatient rotations from July 1, 2015 through June 30, 2016. Resident schedules are broken into 4-week “blocks” with 13 blocks total through the 52-week study period.

Residents were classified according to their role on an inpatient rotation. Junior front-line clinicians (Jr FLCs) were defined as first-year residents serving in a front-line ordering clinician role. Senior front-line clinicians (Sr FLCs) were defined as senior residents (postgraduate years 2–4) serving in a front-line ordering clinician role. Senior supervisors

(Sr Suprv) were defined as senior residents (postgraduate years 2–4) serving in a supervisory role.

Statistical Analysis

Summary statistics and data visualizations were created for scheduled versus EHR-calculated shift validations including counts, percent overlap, and frequency distributions of start- and end-time differences. Summary statistics of use case calculated shifts include counts and duration split by role, rotation as well as time of year (quarter). All inpatient pediatric rotations were included in this analysis, including pediatric and neonatal intensive care units, oncology, and all pediatric subspecialties as defined by the Pediatrics Residency Program. Rotations were anonymized for analysis. Continuous variables with multiple groups were analyzed with analysis of variance (ANOVA) followed by Tukey’s post-hoc testing to adjust for multiple comparisons. All data processing and analysis were completed in R Studio.³¹

Results

Comparison of EHR Timestamps to Scheduled Shifts

We collected 6.3×10^7 EHR timestamp events across a 12-month period. Shift validation was completed on 1,237 EHR-calculated shifts. From these validation shifts, each EHR-calculated shift was matched to a scheduled shift with no mismatches, meaning that there were no EHR-calculated shifts without a corresponding scheduled shift and no scheduled shifts without an EHR-calculated shift. The percent overlap of time period was $87.9 \pm 0.3\%$ (mean \pm standard error of mean [SEM]; → Fig. 3A). Specifically, physicians functioning in the role of front-line clinician demonstrated the highest overlap (Jr FLC: $88.8 \pm 0.4\%$ [mean \pm SEM], Sr FLC: $88.5 \pm 0.5\%$) compared with physicians in a supervisory role (Sr Suprv: $85.0 \pm 0.9\%$, ANOVA $p < 0.001$, pairwise comparison $p < 0.001$). Shifts with physicians in all roles had median

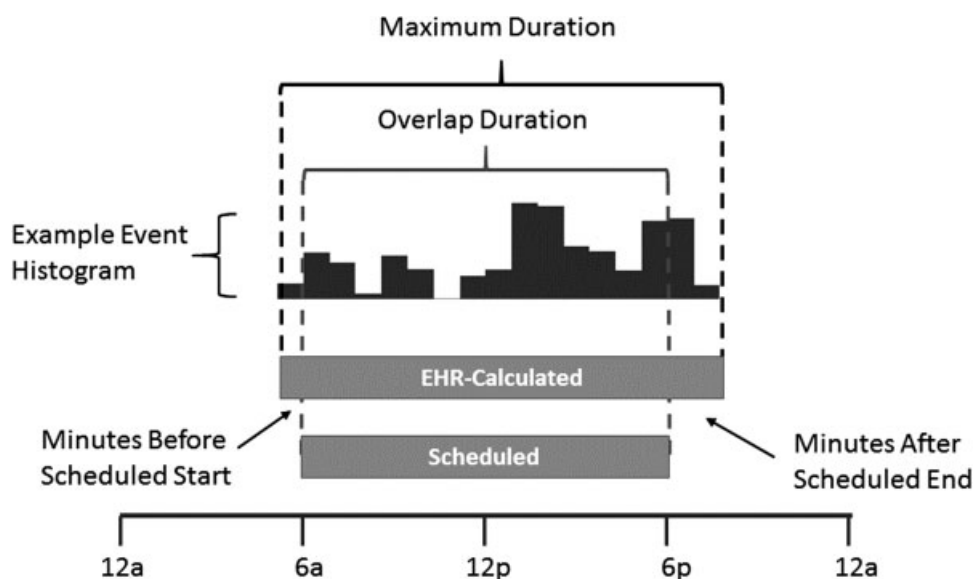


Fig. 2 Shift percent overlap is calculated by dividing the “overlap duration” by the “maximum duration” and reflects the degree of overlap between a scheduled and EHR-calculated shift. An example histogram of EHR event timestamps is shown along with the bounding boxes representing the EHR-calculated shift. The time periods before and after the scheduled shift are indicated by arrows. EHR, electronic health record.

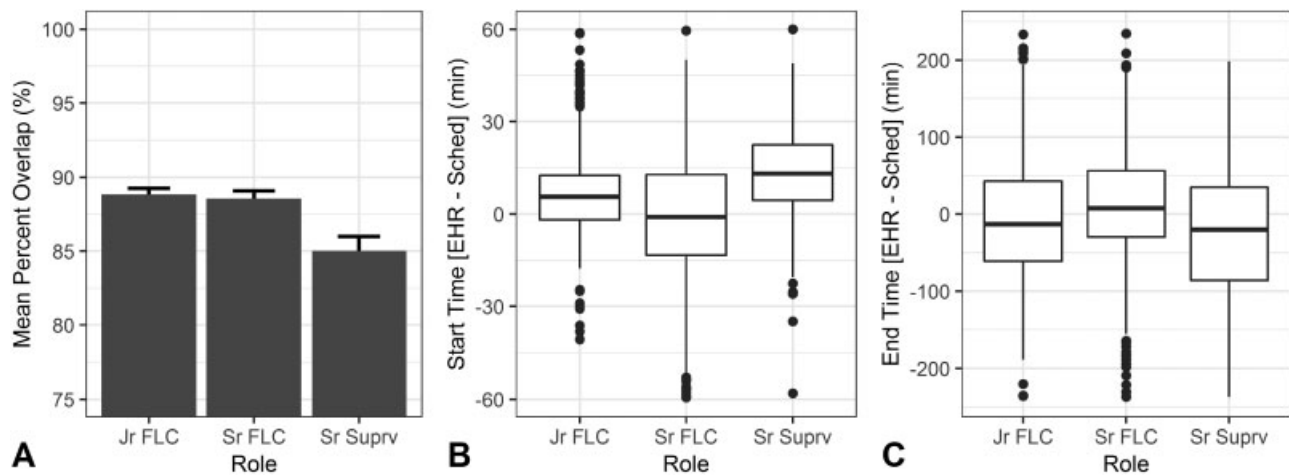


Fig. 3 Validation shifts show excellent agreement between scheduled and EHR-calculated shifts as indicated by percent overlap and start- and end-time differences. (A) Mean percent overlap is high ($\geq 85\%$) across all roles, with significantly higher overlap in both FLC roles (Jr FLC: 88.8 ± 0.4 ; Sr FLC: 88.5 ± 0.5) compared with the Sr Suprv role (mean \pm SEM: 85.0 ± 1.0). (B, C) Box-whisker plots showing the difference between EHR-calculated and scheduled start time (B) and end time (C). A median difference of zero suggests no difference, whereas a positive value indicates the shift started (or ended) later than scheduled and a negative value indicates the shift started (or ended) earlier than scheduled. Boxes represent first through third quartile (25th to 75th percentile) with the median indicated by the thick line inside the box. Whiskers extend to $1.5 \times$ IQR in each direction, with outliers plotted with dots individually. Note that the y-axis limits differ between figures (B) and (C), and are truncated to better show whiskers though outliers continue beyond the limits of the plots. EHR, electronic health record; IQR, interquartile range; SEM, standard error of mean.

EHR-calculated start times close to scheduled start times (Jr FLC: 5.6; Sr FLC: -0.9 ; Sr Suprv: 13.3 minutes; **Fig. 3B**). Additionally, all roles had median EHR-calculated end times close to scheduled end times (Jr FLC: 13.4; Sr FLC: -7.9 ; Sr Suprv: 7.3 minutes; **Fig. 3C**). Shifts with physicians in supervisory roles had EHR-calculated start times 30.9 ± 6.1 minutes (mean \pm SEM) later than scheduled start times and EHR-calculated end times 65.3 ± 10.3 minutes earlier than scheduled, which are significantly greater than those from physicians in FLC roles (ANOVA $p < 0.001$, pairwise comparison $p < 0.0001$); however, mean values are subject to influence by outlier data points.

To better visualize the skew and distribution in shift start and end time differences, we plotted histograms of the difference between EHR-calculated and scheduled start and end times (**Fig. 4**). Histograms were truncated on the x-axis to better visualize the spread around the point of zero difference, with negative values reflecting EHR-calculated shifts starting or ending earlier than scheduled shifts and positive values reflecting EHR-calculated shifts starting or ending later than scheduled shifts. Among differences in shift start time, the Jr FLC role had the narrowest dispersion (standard deviation [SD]: 15.2) compared with those in the Sr FLC (SD: 24.4) and Sr Suprv (SD: 18.0) roles. Differences in

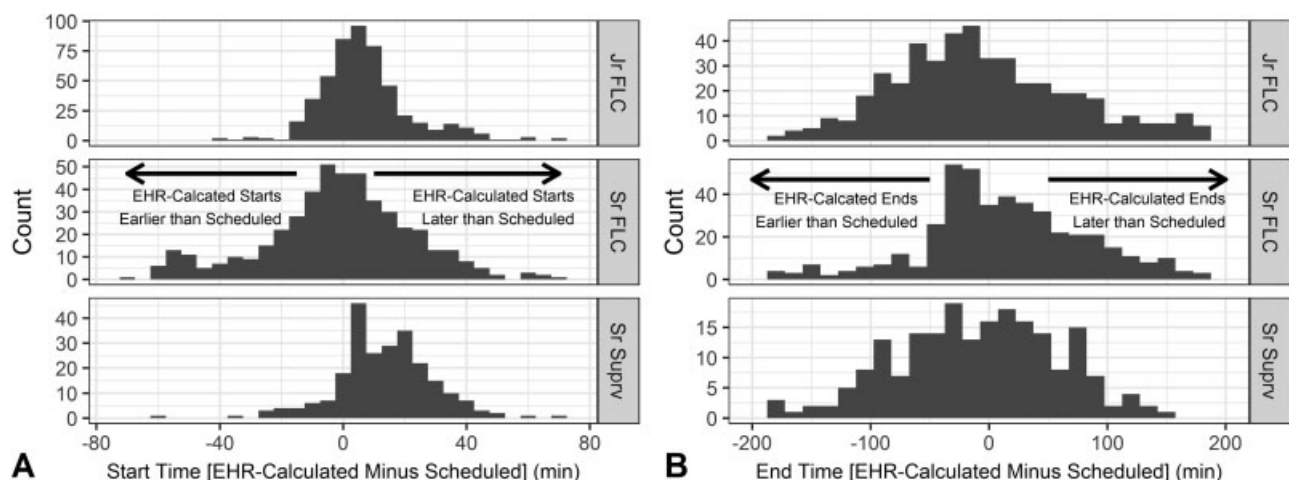


Fig. 4 Histogram distributions of the differences between EHR-calculated and scheduled start and end times show wide variation among validation shifts. (A) Start time difference (EHR-calculated minus scheduled), where negative values indicate an EHR-calculated start time earlier than scheduled. (B) End time difference (EHR-calculated minus scheduled), where negative values indicate an EHR-calculated end time earlier than scheduled. Note that the x-axis limits differ between figures (A) and (B), and are truncated to better show distributions though additional points exist beyond the limits of the plots. EHR, electronic health record.

Table 1 Cohort summary

	Jr FLC	Sr FLC	Sr Suprv
Resident blocks	339	305	127
Resident year (PGY)			
PGY 1	339 (100%)	–	–
PGY 2	–	268 (88%)	59 (46%)
PGY 3	–	37 (12%)	65 (51%)
PGY 4	–	–	3 (3%)
Shift counts			
EHR-calculated shifts	7,096	5,401	2,181

Abbreviations: Jr FLC, junior front-line clinicians; PGY, postgraduate year; Sr FLC, senior front-line clinicians; Sr Suprv, senior supervisors. Note: “Jr FLC” is defined as a first-year resident performing front-line clinical roles; “Sr FLC” is defined as a second-, third-, or fourth-year resident performing front-line clinical roles; and “Sr Suprv” is defined as a second-, third-, or fourth-year resident performing in a supervisory role.

end times had greater dispersion compared with start times, with all three roles having similar SDs (Jr FLC: 78.0; Sr FLC: 73.3; Sr Suprv: 74.3).

Use Case: Pediatric Resident Shifts by Rotation and Role

Our use case cohort consisted of all 144 residents on clinical pediatric inpatient rotations from July 1, 2015 to June 30, 2016. A total of 771 resident inpatient blocks were included in our analysis, distributed by resident role and year as shown in **Table 1**. A total of 14,678 EHR-calculated shifts were captured.

From EHR-calculated shifts we computed average hours worked per rotation block, and compare hours across roles, rotations, and time of year (by quarter). Hours worked were significantly different across roles (ANOVA $p < 0.001$) with Sr FLC blocks having significantly higher number of mean hours worked (273.5 ± 1.7) compared with Jr FLC blocks (241 ± 2.5) and Sr Suprv blocks (253 ± 2.3) (**Fig. 5A**). Hours worked by front-line ordering clinicians (Jr FLC and Sr FLC) also significantly varied by rotation (ANOVA $p < 0.001$), with mean hours worked per block

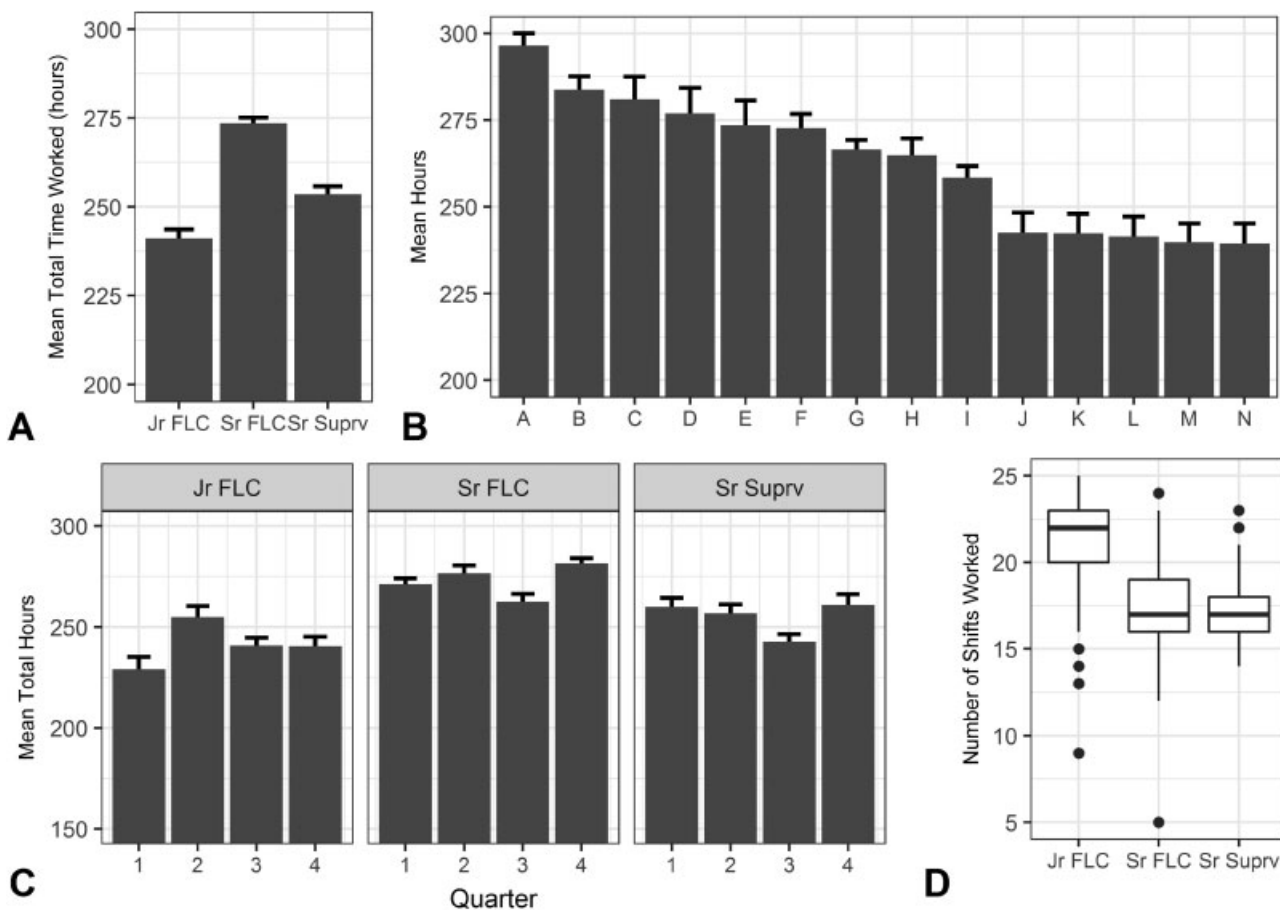


Fig. 5 (A) Mean hours worked varied significantly by role (Jr FLC: 241 ± 2.5 , Sr FLC: 273.5 ± 1.7 , Sr Suprv: 253 ± 2.3 ; $p < 0.001$). (B) Mean hours worked by front-line clinicians also varied by rotation ($p < 0.001$). Rotations include all de-identified Jr FLC and Sr FLC rotations. (C) Mean hours worked did not vary systematically by quarter of the year. (D) Those in the Jr FLC role worked significantly greater number of shifts compared with those in the Sr FLC and Sr Suprv roles ($p < 0.001$). Boxes represent first through third quartile (25th to 75th percentile) with the median indicated by the *thick line* inside the box. Whiskers extend to $1.5 \times$ IQR in each direction, with outliers plotted with *dots* individually. IQR, interquartile range; Jr FLC, junior front-line clinicians; Sr FLC, senior front-line clinicians; Sr Suprv, senior supervisors.

ranging from 239 ± 5.8 to 296 ± 3.4 hours, or an average difference of 14.75 hours per week in a typical 4-week block (→ Fig. 5B). Systematic differences were not observed when stratifying hours worked by quarter (→ Fig. 5C). The number of shifts worked on average was significantly different across roles (ANOVA $p < 0.001$), with Jr FLC blocks having significantly more shifts (21 ± 0.1) compared with Sr FLC (18 ± 0.1) or Sr Suprv (18 ± 0.1) blocks (→ Fig. 5D).

Discussion

Our study is the first to report on the use of EHR timestamps to automatically calculate inpatient provider shifts from EHR timestamps. As our salaried employees are not required to clock in and out, we currently have no gold standard against which to compare, and the best measure of time at work currently is scheduled shifts. Our manual validation of EHR-calculated shifts against these scheduled shifts for trainee physicians working in different roles found excellent agreement for providers in front-line provider roles. Our shift agreement was somewhat lower among providers in supervisory roles, which likely reflects fewer interactions with the EHR at the beginning and end of their shift. In our inpatient workflow it is not uncommon for senior supervisory residents to interact minimally with the EHR while leading rounds, supervising patient admissions, and following up on other patient-care tasks. The shift agreement may be lower among physicians in supervisory roles without EHR interactions “bookending” their shifts. Alternatively, actual shift times worked may differ from the *a priori* defined shift start and end times in our validation cohort.

Furthermore, we have demonstrated a use case of our automated system by extracting trainee-physician inpatient shifts and showing differences across rotations and roles. As these EHR-reported shifts can provide details about hours and shifts worked as well as start and end times, this use case would be of interest to those in charge of training programs, as well as other hospital administrators working with shift-based employees who interact frequently with the EHR (e.g., nurses or hospitalist physicians). The recording of shifts is considered a “nonclinical clerical task” which can be burdensome and can increase burnout.² Additionally, we observed wide variation in hours worked by rotation. Training program directors and course directors can use this objective information to make informed decisions about trainee rotations. Importantly, we show actual times for providers beginning and ending shifts where previously this information was not available. Objective data about providers consistently working later than scheduled shifts, for example, could prompt a review of the processes causing this overtime and improve provider satisfaction.

There are numerous practical applications for these calculated shifts. These data could be presented in summary reports to hospital or program leadership, or distributed to individual providers for their own consumption. In the ideal state, we envision using this technology to develop a “dashboard” of up-to-date shift times across an organization based on EHR usage. Access to these summary data must include

discussions of data privacy, including patient-level information which was not needed for shift calculations, as well as provider-level information. Although the Health Insurance Portability and Accountability Act (HIPAA) mandates mechanisms to “record and examine activity in information systems,”³² care must be taken to prevent unauthorized secondary use of these audit logs. Security restrictions and policies around access log data, as well as employee agreements or consent to aggregate these data, should be included as part of an implementation plan. This research study was Institutional Review Board approved and providers were de-identified to limit individual exposure; however, in operationalizing this tool such de-identification would not be possible.

Several groups have used EHR timestamps to estimate shifts for both outpatient and inpatient providers. Gilleland et al retrospectively queried trainee EHR usage in an outpatient clinic based on login and logout times and compared with self-reported EHR usage.²⁴ They did not calculate a full “shift duration” but rather looked at cumulative time on the EHR; however, it would be relatively straightforward to apply our technique to outpatient data and calculate an outpatient “shift.” To our knowledge, the only other application of EHR timestamps to inpatient shift calculation was reported by Shine et al in their study of internal medicine residents rotating through their medical intensive care unit.²⁷ They collected timestamps and pasted these into a custom spreadsheet which calculated shifts by considering trainees “working” if no 6-hour period passed without EHR interaction. In our study, we iteratively improved upon the shift detection algorithm and automated the collection process, while our use case was expanded to include all trainee physicians working on inpatient rotations.

Developing this system required minimal assistance from our IS group, as our institution already archives EHR timestamps in a long-term storage database. All timestamp data are based on “out-of-the-box” functionality in the EHR and required no additional EHR customization. We included all events (i.e., all metric IDs in our EHR) and therefore the presence or absence of key events would not impact this algorithm; however, decreasing the overall number of events would decrease granularity in event timestamps and could impact the algorithm performance or require changes to the heuristics. For example, although minimal temporal granularity is required as the algorithm makes use of a “first” and “last” event to define the shift, if providers engage in tasks away from the EHR for a long period of time (e.g., educational sessions or procedures) it is possible that the actual shift would be split into multiple “calculated” shifts. Overcoming these splits would require heuristics be set appropriately for the workflows of providers. We estimate it would be feasible to migrate this system to another institution with minimal effort, including customizing heuristic values and adjusting inputs based on EHR timestamp formatting. We developed our algorithm in three iterations, making use of “expert knowledge” of trainee clinical workflows including the structure of calls and night shifts among our providers. While we expect that this approach would translate to other workflows and institutions with minimal change, validating

calculated shifts against an existing standard would be important before assuming generalizability. We would recommend an approach similar to ours to develop shift-interval cut points specific to other workflows, including histogram analysis and validation on a set of scheduled or observed shifts.

Our study is the first to describe a methodology which could be used to conduct large-scale, near-real-time automated monitoring of inpatient physician shift durations and shift end times. While these data may be inherently interesting to regulatory bodies and hospital administrators, they also could allow for specific, targeted interventions to support trainees and improve educational environments in training programs. At the program level, specific clinical rotations that are associated with high workload and low educational value could be targeted for structural changes and curricular improvements. For example, the Pediatrics Residency Program at Children's Hospital of Philadelphia was provided feedback on rotation-level data which will influence curricular planning for subsequent years. Additionally on the individual level, trainees who may be struggling with finishing their clinical workload in a timely manner could be identified and assisted by their training program. Importantly, if individuals were aware that such monitoring was in use, there exists the potential for "Hawthorne effect" behavior modification including consciously shifting EHR activity to alter shift calculations. Such behavior could alter provider workflows and should be considered in implementation strategies.

Graduate medical education officials may be interested in this tool as a real-time monitor of trainee duty hours, as EHR-calculated shifts provide bounds on in-hospital EHR activity and therefore cannot easily overestimate hours worked. Self-reporting systems are vulnerable to recall bias and both under- and overreporting,^{30,33} and trainees across many specialties underreport duty hours.^{28,29,34} Lastly, these data may provide objective outcomes for intervention studies to reduce burnout, as well as fill operational needs to monitor the burden of provider shifts. The low-cost, rapidly available, objective data provided by our methodology have the potential to have a direct impact on the work environment across many fields of medicine.

Limitations

Despite the automated nature of our study and the construct validity of our approach, there are limitations which must be acknowledged. First, our algorithm collected events logged only from in-hospital computers. Because of this, EHR-calculated shifts could very rarely overestimate hours worked (e.g., if the provider was reviewing patient charts for research purposes while physically at the hospital), but could more easily underestimate hours if the provider was working without interacting with the EHR at the beginning or end of their shift or working remotely. This would also occur in the event of long periods away from the EHR such as for conferences, educational sessions, paper-based handoffs, or during periods of EHR downtime. Second, we randomly sampled ~10% of providers to validate against scheduled

shifts; however, this small sampling could bias against infrequent scheduling cases. A larger sampling size would limit this bias. Additionally, we did not include "work from home" specifically because we were interested in inpatient provider shifts, but the addition of this information from timestamp data is relevant as work-home conflicts contribute to burnout.^{2,35}

This study relied on retrospective data analysis and therefore lacked the direct observation of physicians as the gold standard for EHR-calculated shift measurement. Shift validation compared EHR-calculated shifts to scheduled shifts, but did not include direct observation which would be costly and time-consuming. Scheduled shift start and end times were *a priori* defined based on residency program standards, and were not modified once defined. While modification of these scheduled shift start and end times might have led to higher percent time overlap, the intent was not to match perfectly as there is inherent variability in when trainees arrive and leave shifts but rather to show construct validity in our algorithm and heuristics. The expansion of this work to validate against a true gold standard could provide evidence for both this algorithm's effectiveness as well as the challenges in using scheduled shifts or self-reported duty hours for monitoring workload. Finally, we validated our algorithms based on assumptions for trainee-physician workflows in a pediatric residency program at our institution, but these assumptions might vary in other specialties or at other institutions.

Conclusion

Physician burnout is tied to shift duration, clinical workload, and work-life integration. Additionally, excessive physician work hours are associated with worse clinical performance,¹³ increased medical errors,¹⁴ and may contribute to burnout. Secondary use of EHR timestamp data can accurately calculate inpatient provider shifts, and this automated shift calculation can provide objective outcome data for intervention studies as well as monitor shift burden. This algorithm minimizes the clerical burden of data collection, eliminates recall bias, and avoids disincentives for reporting hours. Possible uses include automated shift extraction for trainees and administrators as well as providing objective outcomes in targeted intervention trials to reduce time at work and burnout. Future studies to perform validation across institutions with different workflows will determine the scalability of this approach.

Clinical Relevance Statement

EHR timestamp data can accurately and automatically calculate inpatient provider shifts, measures which could be of interest to hospital administrators, regulatory bodies, and graduate medical education officials. Application of these techniques to pediatric residents demonstrates criterion validity compared with scheduled shifts. Future studies will generalize heuristics to other institutions with different inpatient workflows.

Multiple Choice Questions

1. The chief nursing informatics officer (CNIO) at your institution is interested in knowing what proportion of bedside nurses are working beyond their scheduled shift time. You suggest calculating shifts using EHR access log data. Which of the following statements is true?
 - a. EHR access log data will show the same type of information at every institution, regardless of EHR implementation and settings.
 - b. If EHR access log events are limited to in-hospital workstations, it is very unlikely that the algorithm will overestimate work hours beyond scheduled shifts.
 - c. EHR access log information can only be reliably extracted for physicians.
 - d. Because your CNIO is a user of the EHR, and there are no privacy concerns with respect to these data, she or he should have complete access to the EHR access log data.

Correct Answer: The correct answer is option b, which states that it is very unlikely that the algorithm will overestimate work hours provided the events of interest are limited to in-hospital workstations. Put another way, if an event is logged placing a user on a hospital workstation at a given time, that user must have still been in the hospital at that time, either on a shift or reviewing patient charts for other purposes while physically in the hospital (a less frequent occurrence). Choice a is incorrect because different EHR vendors will store different access log events and even within the same EHR vendor, settings dictate the level of logging. Choice c is incorrect because all users of the system are logged, not just physicians. And choice d is incorrect because there are security and privacy issues surrounding the storage and retrieval of these data; care should be taken when assigning permission for their use.

2. The graduate medical education office would like to report on the duty-hour compliance of its internal medicine residency program. Your designated institutional official (DIO) asks you to duplicate the work described in this article for the purpose of reporting on trainee duty hours. What must be done to ensure this implementation is accurate and correct?
 - a. The algorithm heuristic cut points can be applied to all residency programs without any additional validation, as they were criterion valid compared with scheduled shifts at a single pediatrics residency program.
 - b. No additional modifications are necessary as long as the algorithm is applied only to supervising senior residents, on whom the algorithm has been shown most effective.
 - c. Analysis of frequency distributions of event timestamp intervals from known shifts, as well as workflow analysis of trainee shifts, will help determine heuristic cut points.
 - d. Validation should be undertaken to compare algorithm with scheduled shifts for surgical residents as they are most closely related to the initial study population.

Correct Answer: The correct answer is option c, which describes several of the steps suggested in this article to apply this algorithm to another institution's workflow. Both choices a and b suggest that no additional modifications are necessary, which is likely not true as the heuristic cut points were specifically identified and defined on our own institution's dataset. Additionally, choice d incorrectly suggests that surgical residents are closest to our study population of inpatient pediatrics residents.

Protection of Human and Animal Subjects

This study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects, and was reviewed and approved by the Children's Hospital of Philadelphia Institutional Review Board.

Conflict of Interest

None declared.

Acknowledgments

The authors would like to thank the leadership of the Pediatric Residency Program for their support and assistance, especially the program director Dr Lisa Zaoutis for her thoughtful review of this work; Dr Joseph St Geme, the chair of the Department of Pediatrics, for his advice and support; Dr Anthony Luberti, the clinical informatics fellowship director, for his advice and leadership; and the Enterprise Analytics and Reporting team for providing access and support to the IBM DB2 Hadoop database. Most importantly, we thank the 2015–2016 CHOP pediatrics residents for all that they do in the care of their patients and each other.

References

- 1 Shanafelt TD, Dyrbye LN, West CP. Addressing physician burnout: the way forward. *JAMA* 2017;317(09):901–902
- 2 West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. *J Intern Med* 2018;283(06):516–529
- 3 Roberts DL, Shanafelt TD, Dyrbye LN, West CP. A national comparison of burnout and work-life balance among internal medicine hospitalists and outpatient general internists. *J Hosp Med* 2014;9(03):176–181
- 4 Carayon P, Wetterneck TB, Alyousef B, et al. Impact of electronic health record technology on the work and workflow of physicians in the intensive care unit. *Int J Med Inform* 2015;84(08):578–594
- 5 Babbott S, Manwell LB, Brown R, et al. Electronic medical records and physician stress in primary care: results from the MEMO Study. *J Am Med Inform Assoc* 2014;21(e1):e100–e106
- 6 Coleman M, Dexter D, Nankivil N. Factors affecting physician satisfaction and Wisconsin Medical Society strategies to drive change. *WMJ* 2015;114(04):135–142
- 7 Gregory ME, Russo E, Singh H. Electronic health record alert-related workload as a predictor of burnout in primary care providers. *Appl Clin Inform* 2017;8(03):686–697
- 8 Shanafelt TD, Dyrbye LN, Sinsky C, et al. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clin Proc* 2016;91(07):836–848

- 9 Desai SV, Asch DA, Bellini LM, et al; iCOMPARE Research Group. Education outcomes in a duty-hour flexibility trial in internal medicine. *N Engl J Med* 2018;378(16):1494–1508
- 10 Destino LA, Valentine M, Sheikhi FH, Starmer AJ, Landrigan CP, Sanders L. Inpatient hospital factors and resident time with patients and families. *Pediatrics* 2017;139(05):e20163011
- 11 Fletcher KE, Visotcky AM, Slagle JM, Tarima S, Weinger MB, Schapira MM. The composition of intern work while on call. *J Gen Intern Med* 2012;27(11):1432–1437
- 12 Neri PM, Redden L, Poole S, et al. Emergency medicine resident physicians' perceptions of electronic documentation and workflow: a mixed methods study. *Appl Clin Inform* 2015;6(01):27–41
- 13 Gordon JA, Alexander EK, Lockley SW, et al; Harvard Work Hours, Health, and Safety Group (Boston, Massachusetts). Does simulator-based clinical performance correlate with actual hospital behavior? The effect of extended work hours on patient care provided by medical interns. *Acad Med* 2010;85(10):1583–1588
- 14 Landrigan CP, Rothschild JM, Cronin JW, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med* 2004;351(18):1838–1848
- 15 Lopetegui M, Yen PY, Lai A, Jeffries J, Embi P, Payne P. Time motion studies in healthcare: what are we talking about? *J Biomed Inform* 2014;49:292–299
- 16 Sinsky C, Colligan L, Li L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med* 2016;165(11):753–760
- 17 Zheng K, Guo MH, Hanauer DA. Using the time and motion method to study clinical work processes and workflow: methodological inconsistencies and a call for standardized research. *J Am Med Inform Assoc* 2011;18(05):704–710
- 18 Chen L, Guo U, Illipparambil LC, et al. Racing against the clock: internal medicine residents' time spent on electronic health records. *J Grad Med Educ* 2016;8(01):39–44
- 19 Wu DTY, Smart N, Ciemins EL, Lanham HJ, Lindberg C, Zheng K. Using EHR audit trail logs to analyze clinical workflow: A case study from community-based ambulatory clinics. *AMIA Annu Symp Proc* 2018;2017:1820–1827
- 20 Arndt BG, Beasley JW, Watkinson MD, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med* 2017;15(05):419–426
- 21 Hribar MR, Read-Brown S, Goldstein IH, et al. Secondary use of electronic health record data for clinical workflow analysis. *J Am Med Inform Assoc* 2018;25(01):40–46
- 22 Hirsch AG, Jones JB, Lerch VR, et al. The electronic health record audit file: the patient is waiting. *J Am Med Inform Assoc* 2017;24(e1):e28–e34
- 23 Goldstein IH, Hribar MR, Read-Brown S, Chiang MF. Association of the presence of trainees with outpatient appointment times in an ophthalmology clinic. *JAMA Ophthalmol* 2018;136(01):20–26
- 24 Gilleland M, Komis K, Chawla S, Fernandez S, Fishman M, Adams M. Resident duty hours in the outpatient electronic health record era: inaccuracies and implications. *J Grad Med Educ* 2014;6(01):151–154
- 25 Vawdrey DK, Wilcox LG, Collins S, et al. Awareness of the care team in electronic health records. *Appl Clin Inform* 2011;2(04):395–405
- 26 Hom J, Richman I, Chen JH, Singh B, Crump C, Chi J. Fulfilling outpatient medicine responsibilities during internal medicine residency: a quantitative study of housestaff participation with between visit tasks. *BMC Med Educ* 2016;16:139
- 27 Shine D, Pearlman E, Watkins B. Measuring resident hours by tracking interactions with the computerized record. *Am J Med* 2010;123(03):286–290
- 28 Byrne JM, Loo LK, Giang DW. Duty hour reporting: conflicting values in professionalism. *J Grad Med Educ* 2015;7(03):395–400
- 29 Carpenter RO, Spooner J, Arbogast PG, Tarpley JL, Griffin MR, Lomis KD. Work hours restrictions as an ethical dilemma for residents: a descriptive survey of violation types and frequency. *Curr Surg* 2006;63(06):448–455
- 30 Saunders DL, Kehoe KC, Rinehart VH, Berg BW. Self-reporting of internal medicine house staff work hours. *Hawaii Med J* 2005;64(01):14–16
- 31 R. Core Team. *R: A Language and Environment for Statistical Computing* [program]. Vienna, Austria: R Foundation for Statistical Computing; 2017
- 32 Health Insurance Portability and Accountability Act. *Security Standards for the Protection of Electronic Protected Health Information*. 45 C.F.R. 2003
- 33 Gonzalo JD, Yang JJ, Ngo L, Clark A, Reynolds EE, Herzig SJ. Accuracy of residents' retrospective perceptions of 16-hour call admitting shift compliance and characteristics. *J Grad Med Educ* 2013;5(04):630–633
- 34 Bennett CL, Finch A, Vuong K, McDonald D, Rennie S. Surgical resident duty hours. *N Engl J Med* 2016;374(24):2399–2401
- 35 Shanafelt TD, Hasan O, Dyrbye LN, et al. Changes in burnout and satisfaction with work-life balance in physicians and the general US working population between 2011 and 2014. *Mayo Clin Proc* 2015;90(12):1600–1613