Emergency Physician Workload

Grant Award Number: R03-HS024801

Principal Investigator: Kathryn M. Kellogg, MD, MPH

Team: Allan Fong, MS

Raj Ratwani, PhD

Rollin J. Fairbanks, MD, MS

Amy Will

Tracy Kim

Organization: MedStar Health

MedStar Institute for Innovation

National Center for Human Factors in Healthcare

Dates of Project: 9/01/16-8/29/18

Federal Project Officer: Denise Burgess

Acknowledgement of Agency Support: Agency for Healthcare Research and Quality

Contents

Purpose:

The emergency department (ED) is a complex and dynamic clinical environment. Emergency physicians face rapidly changing patient volumes, patients with emergent conditions, and task interruptions that can result in periods of incredibly high workload and stress. Increased workload and stress can negatively impact physician memory, perception, and reasoning, resulting in potential patient safety hazards. Safety hazards that may be a result of increased workload include ordering the wrong medication, ordering the wrong lab or imaging study, or selecting the wrong patient from the electronic health record.

The purpose of this study was to develop measures of physician workload by correlating observational data of physician workflow processes with physiological data gathered from physicians while they were performing their clinical activities. Developing reliable measures of physician workload can then serve as an indicator as to when physician workload may need to be better balanced across the care team to maximize effectiveness and reduce the likelihood of patient safety hazards.

Scope:

Specific Aim 1:

Define stress-associated factors in the emergency department using physiologic response data and ethnographic in-context observational data.

Output from Specific Aim 1:

A validated application to concurrently record heart rate and activity data to be used during observational studies

A list of contextual stress-associated factors in the emergency department, derived through observation correlated with physician physiologic response

Specific Aim 2:

Validate the contextual stress-associated factors identified in Specific Aim 1 through semi-structured interviews with an independent group of practicing emergency physicians. We hypothesize that physicians will identify previously reported stress-inducing factors and will validate the factors identified in Specific Aim 1.

Output from Specific Aim 2:

A taxonomy of stress-associated factors for physicians in the emergency department

Methods:

Specific Aim 1:

In this aim, we developed and validated a heart rate (HR) monitor that could integrate a live HR data stream into an application used to track in real time the activities being performed by physicians in the clinical setting. We then performed observations of physicians working in the ED and performed analyses of their HR and heart rate variability (HRV) to identify those activities associated with the highest HR or greatest HRV.

The application (app) used for this Specific Aim, called TaskTracker, was developed at our Center and is shown in Figure 1, below. This app is built with HTML5 and runs on a Windows tablet. It allows an observer to categorize the general work processes of the physician, such as time on the computer, time with the patient, time on the phone, etc. The app also provides a time stamp to record when tasks are started and stopped.

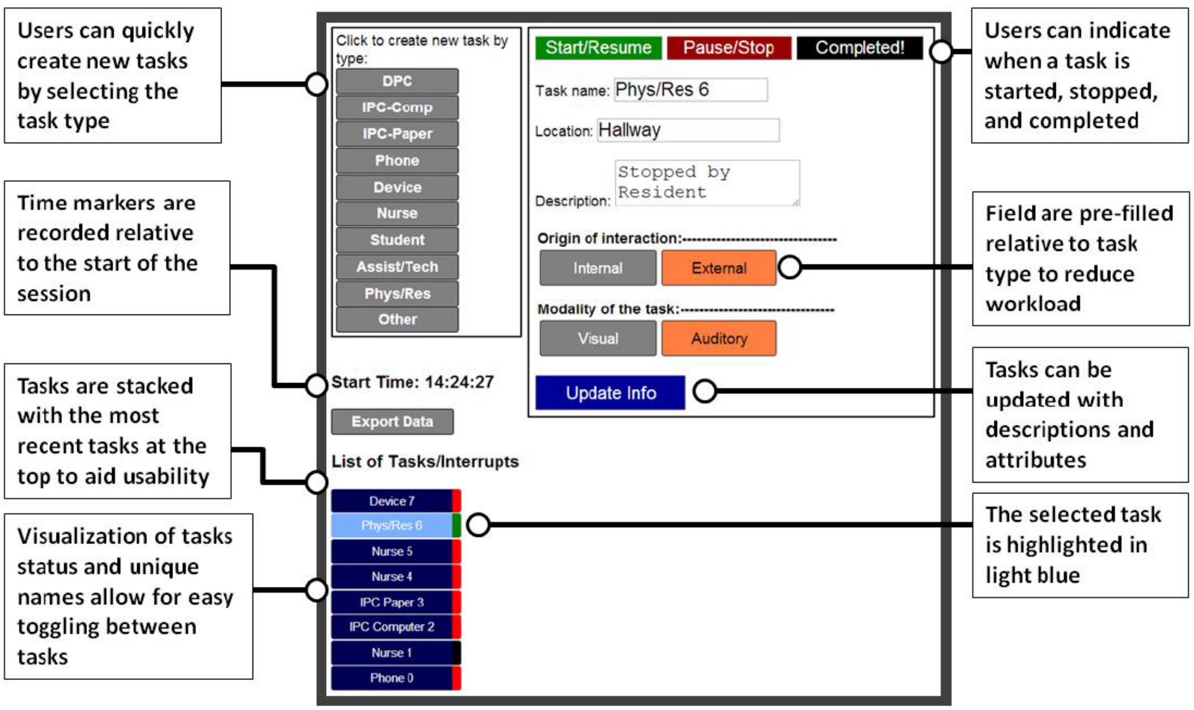

Figure 1. TaskTracker.

Heart rate sensor validation and data integration:

Although we performed pilot observations using a commercially available sensor, multiple difficulties became clear as we looked to scale the study to include more participants. The multiple commercially available sensors we tested did not allow us to stream and access electrocardiogram (ECG) data in real time during observations. It was also difficult to acquire actual ECG data versus reported heart rate, and the reliability of that heart rate reporting was unclear. Additionally, as these were commercial products, our team was unable to modify the proprietary algorithms to fit our specific needs, and we were unable to integrate the heart rate feed into our TaskTracker application for accurately integrated timing of the data. Commercially available 3-lead ECG sensors were available that may have more appropriately fit our research needs with open-source data, but these were cost prohibitive. These limitations made the affordable commercially available sensors not scalable for this project. Given this limitation, our team initiated development and validation of an ECG sensor that would be modifiable, would be comfortable for a wide range of participants, and could be integrated into the TaskTracker application.

The ECG sensors used in our sensor are produced by Bitalino (www.bitalino.com), which can sample up to 1000 samples per second. Figure 2 diagrams the system we developed to integrate the ECG sensor with the observational TaskTracker tool, which runs on a Windows tablet. The tablet operating system first initiates a server on python, which connects to the ECG sensor via Bluetooth. The ECG data stream is batched and processed in python and recorded to a data log.

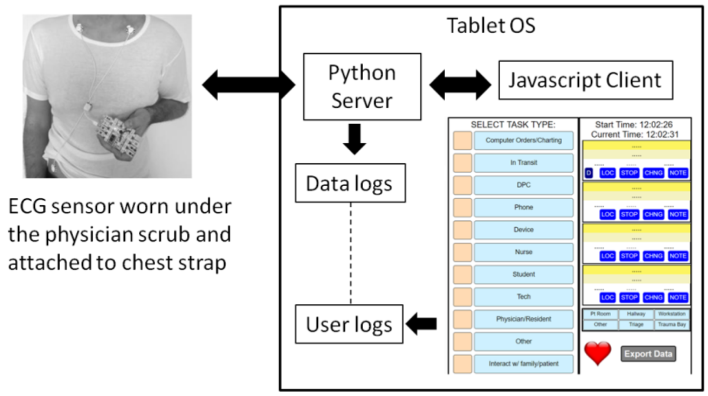

Figure 2. Integrated ECG and TaskTracker app.

At the same time, a real-time heartbeat peak-finder algorithm was developed. Every time a heartbeat is identified by the python algorithm, it is passed to the Javascript client, which then animates a beating heart to provide visual feedback to the observer. In addition, the Javascript client sends an update message intermittently to the server to ensure consistent timestamps. The observer then uses TaskTracker to record physician tasks. Heartbeats are displayed in real time. In addition, user collected observational data is logged using the same timing, allowing for post-hoc data analysis.

The ECG sensor was validated in a laboratory setting. Participants were recruited to perform a series of different tasks in 2-minute intervals (sit, stand stationary, walk at 1 mph on a treadmill, walk at 2 mph on a treadmill, stand stationary) while connected to our custom-developed ECG system and a standard pulse oximeter (FDA approved, error of +/- 2 bpm). The pulse oximeter was used as the gold standard by which the ECG monitor was evaluated. These tasks were selected to mimic common body positions and movements of emergency physicians outside the patient room. Pearson product-moment correlation and standard error of estimates were calculated.

For analysis of ECG output, ECG peaks were identified for each valid 10-second time period. Extraneous peaks were removed with a smoothing window. Next, the RR intervals between heart beats were calculated and outliers, such as RR intervals corresponding to heart rates less than 40 bpm or greater than 200 bpm, were removed. Additionally, HR was recorded for valid time windows when at least 80% of RR intervals were within a 25% error range. Data were then processed by synchronizing the 10-second sections with the task data. By breaking this into 10-second segments, we were able to identify changes in HR during even very short tasks. For instance, during a 40-second computer task segment below, HR increased and then decreased during the segment (Figure 3).

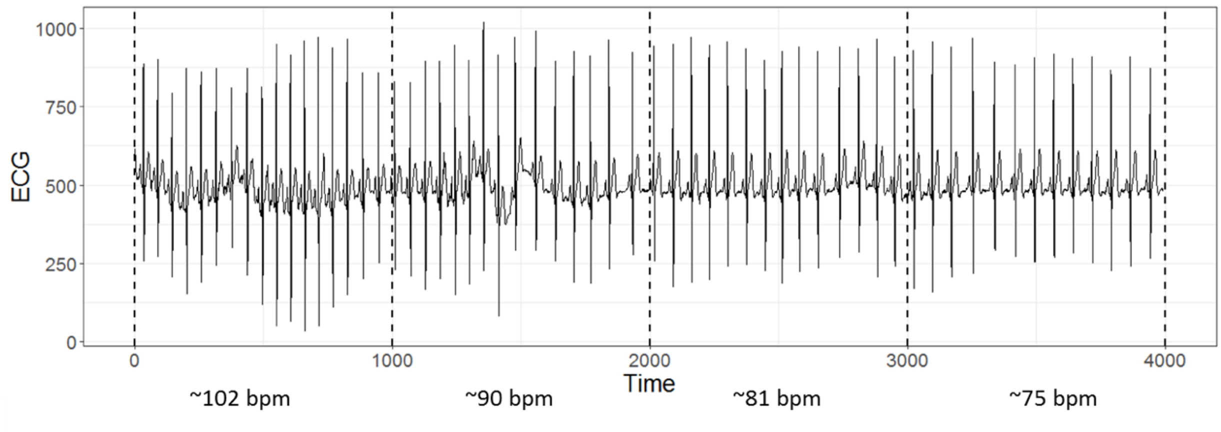

Figure 3. ECG tracing during a computer task.

The average HR for each participant and across participants was calculated. Given differences between individuals, each participant's HR data was analyzed independently. Each 10-second segment was compared to the participant's average HR and placed into categories, below.

- Zone 1: HR_i is less than $\overline{HR}$ for the participant: $HR_i < \overline{HR}$
- Zone 2: HR_i is between $\overline{HR}$ and one positive standard deviation, SD(HR), of the participant's $\overline{HR}$: $\overline{HR} <= HR_i < \overline{HR} + SD(HR)$
- Zone 3: HR_i is between the one and two positive standard deviation of the participant's $\overline{HR}$: $(\overline{HR} + SD(HR)) <= HR_i < (\overline{HR} + 2SD(HR))$
- Zone 4: HR_i is greater then or equal to two positive standard deviation of the participant's $\overline{HR}$: $(\overline{HR} + 2SD(HR)) <= HR_i$

The time spent in these zones was compared to task data for each participant, as illustrated in Figure 4.

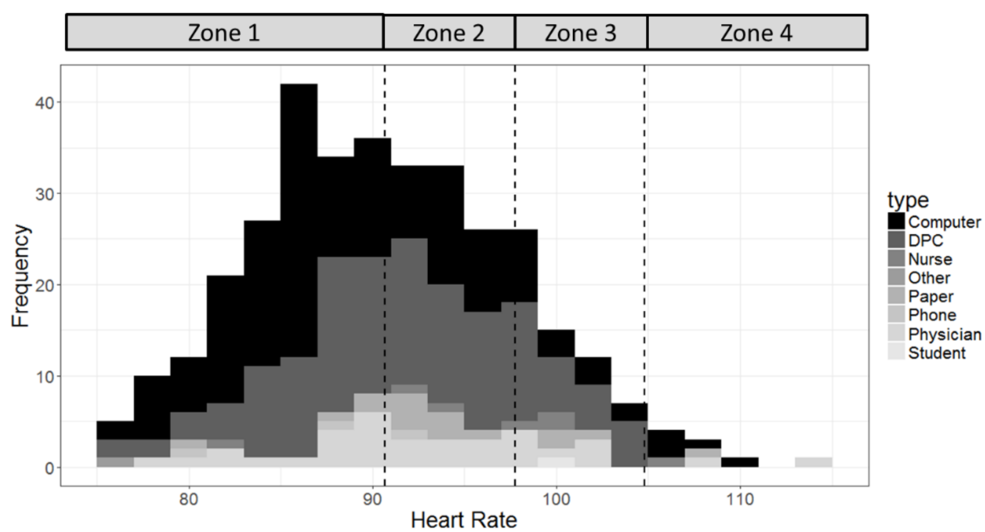

Figure 4. HR zones and tasks performed by one participant.

Then, we calculated a ration of occurrence to determine if certain tasks were more likely to be associated with the higher HR zone.

<u>Observational Data Collection</u>: Emergency medicine attending physicians and residents at two inner city institutions were recruited via email and word of mouth for participation. A researcher observed emergency physicians for 2 hours per session to document the work tasks and workflow processes of the physician. The observer stood at a distance of 5 to 10 feet from the physician and used a web based application on a tablet to document the work processes as one of 11 tasks. The app was updated with these tasks according to previous experience with task observation in the ED (Table 1). Observers did not enter patient rooms. Two observers completed the data collection. A pilot period of observations wherein both observers recorded data simultaneously was performed, and a kappa was calculated using both sets of data to establish reliability. Descriptive statistics around time spent on different tasks were calculated. Detailed analysis of the correlation of tasks and HR is ongoing.

Table 1. Tasks included in the TaskTracker app for ED physician observations.

Task Type	Definition
Computer	Participant working on a computer
Device	Participant using any device other than a phone or computer (e.g., ultrasound machine, EKG)
Direct Patient Care	Participant seeing a patient
Family	Participant speaking to a patient's family member
Nurse	Participant speaking or interacting with a nurse
Other	Participant interacting with any other staff or object, or completing tasks not covered by the previous categories (e.g. speaking to a secretary, eating)
Paper	Participant working with paper notes, charts, etc.
Phone	Participant using a phone
Physician	Participant speaking or interacting with another physician
Student	Participant speaking or interacting with a student
Technician	Participant speaking or interacting with a technician or assistant

NASA-Task Load index (TLX)

Twice during the observation, at the start and end, the participant was asked to complete the NASA-TLX. This six-question instrument measures self-reported workload in the following categories: mental demand, physical demand, temporal demand, performance, effort, and frustration. Differences between attendings' and residents' scores, as well as the change in scores over the observation period, were compared using Student's t-test.

Specific Aim 2

Semi-structured interviews:

A convenience sample of emergency attending physicians and residents was recruited to participate in semi-structured interviews. An interview guide (below) was developed using preliminary findings from SA1 as well as from literature data regarding stress and emergency medicine. These interviews were conducted by a research assistant, and data were collected using field notes. Field notes were collated, and a qualitative analysis was conducted to understand emerging themes from the data.

Interview Questions

1. What is your role?
   a. Attending
   b. Resident
2. How long have you been in your current role?
3. What activities or hobbies do you currently have to relieve stress?
4. Would you characterize your clinical work as stressful? Why or why not?
5. Are there factors about the ED environment that you think contribute to your stress level?
   a. What are they
   b. Has this changed over time in your practice?
6. Do you have strategies for managing stress at work?
   a. What are they?
7. Did your residency training teach any formal training on managing stress at work? Or on managing interruptions or multitasking?
8. How often are you interrupted in your clinical environment?
9. How do you refocus yourself after you have been interrupted during a task in your clinical environment?
10. How often do you find yourself multitasking in your clinical environment?
    a. Are there certain tasks that you multitask more than others?
    b. If yes, why?
11. Please rank the following tasks in order from most stressful to least stressful.
    a. Computer orders
    b. Direct patient care
    c. Paperwork
    d. Phone calls
    e. Interacting with other physicians
    f. Interacting with nurses
    g. Moving from task to task (transitioning)
    h. Interacting with students
    i. Interacting with technicians
12. Please rank the following tasks in order from least amount of time spent to most

a. Computer orders
b. Direct patient care
c. Paperwork
d. Phone calls
e. Interacting with other physicians
f. Interacting with nurses
g. Moving from task to task (transitioning)
h. Interacting with students
i. Interacting with technicians

Results:

Specific Aim 1

Heart rate & task tracking

Validation studies of the ECG sensor resulted in a significant Pearson product-moment correlation (r) between the pulse oximeter and the new ECG sensor, as calculated every 10 seconds: $r=0.94$, $t=14$, $p<0.001$. The standard error of estimate (SEE) also was low, at 3.4 bpm.

A total of 24 physicians (12 attending physicians and 12 residents) completed the 2-hour observations, resulting in 2936 minutes of task tracking and HR data. During the first 2-hour pilot observation, both observers recorded task-tracking data simultaneously. For this session, there was an intra-class correlation of 0.998 and a kappa coefficient of 0.91 between the observers. Given these results, the subsequent observations were completed by one of these two observers.

A total of 70% (SD 24%) of all the 10-second HR intervals was considered valid and included in the analysis. Residents generally showed higher average HR for the observation period and for each task type (Figures 5 and 6).

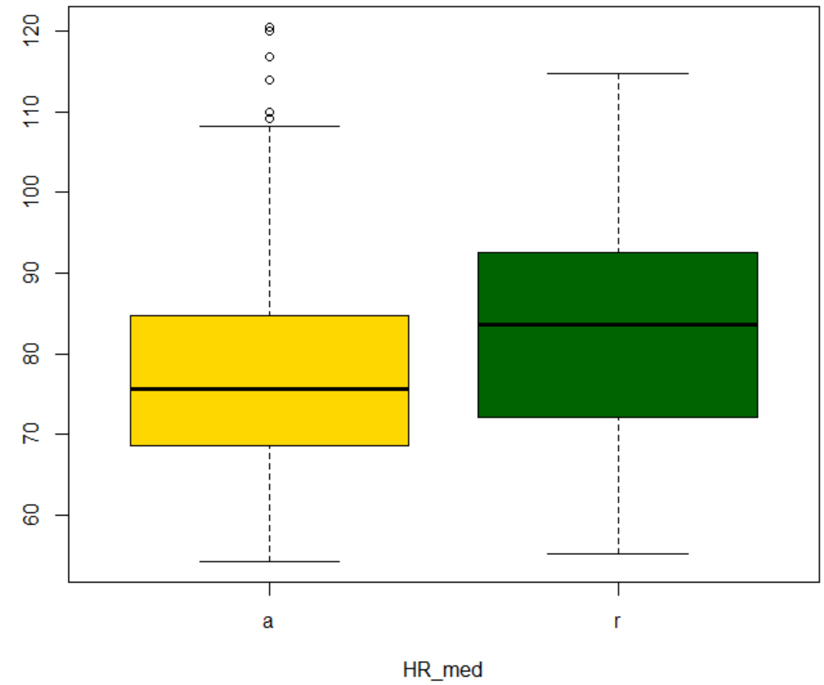

Figure 5. Average HR for attending physicians (yellow) and resident physicians (green).

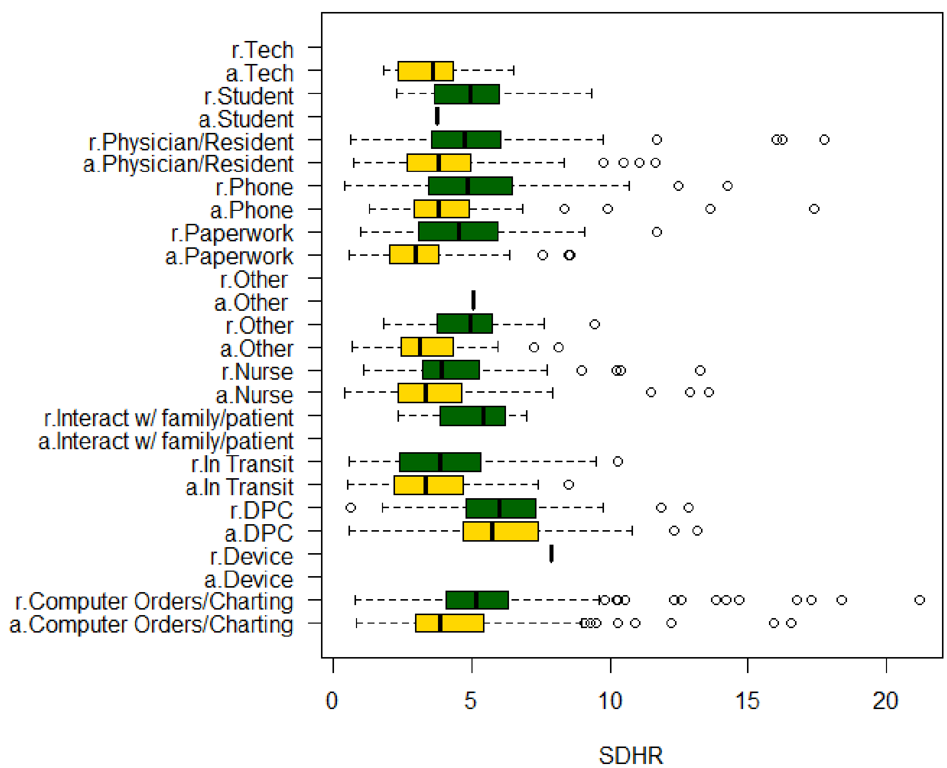

Figure 6. Standard deviation in HR for each task type for attending physicians (yellow) and resident physicians (green).

There was a significant association ($p<0.001$) between the average ratio of occurrence for the HR zones in the computer and direct patient care task categories. In the direct patient care task segments, Zone 3 and 4 segments were more prevalent at the beginning and end of the patient encounter.

Table 2. Ratio of occurrence for HR zones in each task category.

Task Type	Zone 1	Zone 2	Zone 3	Zone 4
Computer	0.43	0.30	0.21	0.13
Device	0	0	0	0
Direct Patient Care	0.39	0.51	0.56	0.57
Family	0	0	0.01	0.01
Nurse	0.02	0.03	0.05	0.05
Other	0.02	0.02	0.03	0.04
Paper	0.01	0.02	0.01	0.01
Phone	0.03	0.02	0.02	0.05
Physician	0.09	0.08	0.09	0.10
Student	0.01	0.01	0.01	0
Technician	0	0	0	0

Further analysis of the HR zones, heart rate variability, and tasks is ongoing in preparation for publication.

NASA-TLX

Overall, post NASA-TLX average scores were higher for all participants, but this difference was not significant. Residents had higher NASA-TLX scores on average in both the pre and post surveys (Figure 7).

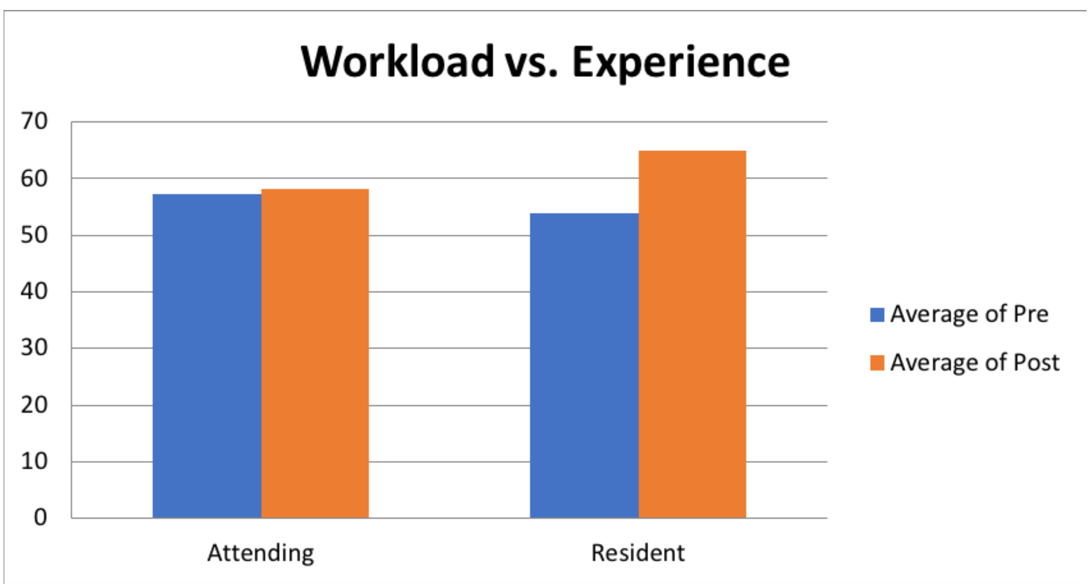

Figure 7. Average NASA-TLX scores reported by participants.

Participants had a wide range of reported pre and post NASA-TLX scores, with no significant difference in the pre and post scores ($p=0.64$) but a general trend toward higher scores in the post survey (Figure 8).

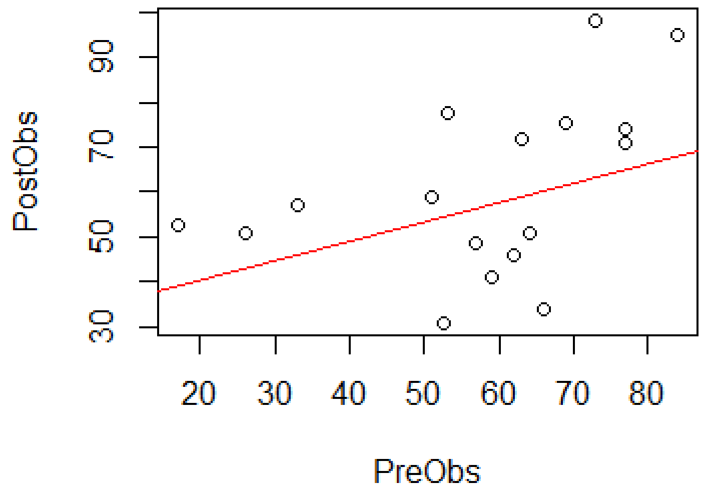

Figure 8. Pre- and Post NASA-TLX scores reported by participants.

We have begun to examine NASA-TLX scores and the time of day when the observation was conducted. There was a trend toward a greater increase in workload scores at the end of a day shift observation, but this was not significant (Figure 9).

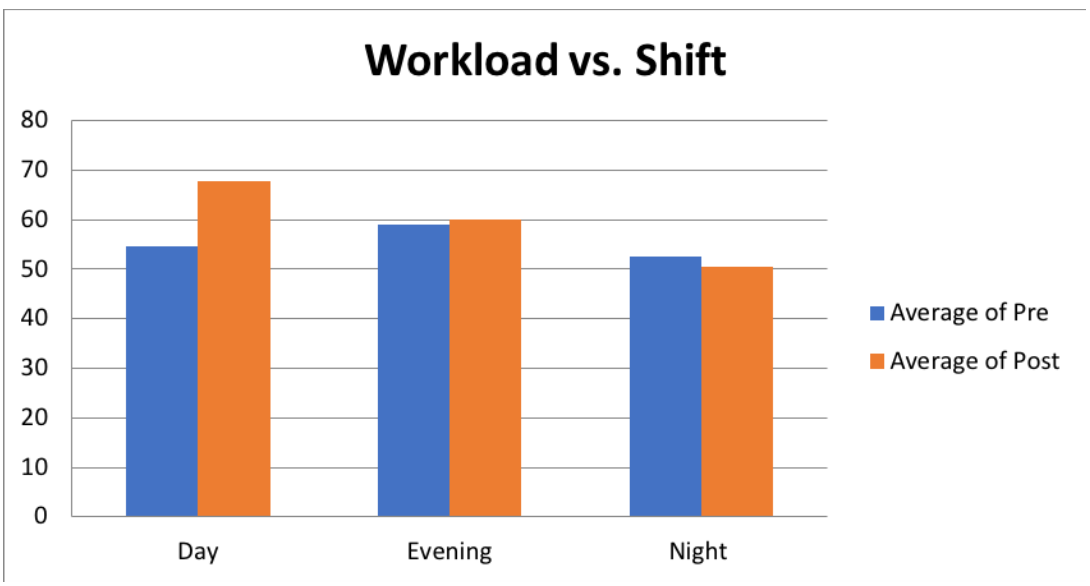

Figure 9. NASA-TLX scores by time of shift.

Further analysis of the NASA-TLX scores in relation to the association of reported scores/change in score with the percentage of time spent on tasks associated with higher HR is ongoing.

Specific Aim 2:

Interviews

Sixteen emergency physicians (n=8 attending physicians and n=8 residents) were recruited to participate; 56% of participants were men, and 44% were women. Resident physicians' years of experience ranged from 1.5 to 3 years, and attending physicians' ranged from less than 1 year to 19 years.

Overall, physicians reported 20 different aspects of their clinical work to be most stressful. Overwhelming workload (44%), multitasking (38%), and high patient volume (32%) had the highest combined consensus. By position, half of the attending physicians reported interruptions, teaching, and boarding to be the most stressful aspects, whereas overwhelming workload, multitasking, and critical decision making were identified by slightly more than half of the residents. These aspects were reported to have not changed over participants' careers, with boarding and interruptions described as worsening. However, half the participants expressed that, for some aspects discussed, stress decreased with experience. When asked to identify the top five most stressful tasks or activities during the shift, interacting with difficult consultants (50%), performing procedures (38%), documenting (31%), and coordinating care with an inefficient team (31%) received the highest consensus among physicians. In relation to patient care, patient expectations were reported to be most stressful (69%).

Attendings found communicating with the care team and sharing mental workload with other physicians to be the most effective aspects of the work environment to decrease these stressors, whereas residents found support from an attending physician to be the most effective in addition to sharing mental workload with others. Overall, 81% of physicians reported having personal strategies for self-management, most which were learned from another attending physician or colleague (56%) or personally sought out through an external source (38%). Taking a 5- to 10-minute physical break off the unit once or twice a shift was the most common strategy described.

An average of 22 interruptions per hour was reported to occur during a typical shift by attending physicians (and 10 by residents), and 13 physicians overall found this rate of interruptions to be stressful. In addition, 69% reported they had received no formal training in multitasking or managing interruptions. Most interruptions were described to take place at the workstation and were most stressful if the primary task was more critical than the interruption. Half the physicians (n=3 attendings, n=5 residents) did not have a refocusing strategy. Interacting with the EHR and the phone were reported to be the primary tasks that physicians multitasked due to the low cognitive load required.

List of Publications and Products:

1. Task2Heart: Integrated TaskTracker/ECGs sensor app
2. Fong, A., Kim, T.C., Ratwani, R.M, Kellogg, K.M. Task2Heart: Exploring Heart Rate Differences with Time-Motion Workflow Observations of Emergency Medicine Physicians. J Med Syst (2018) 42: 170.
3. Kim T, Kellogg K, Nare M, Blumenthal J, Fong A. Understanding the noise: Categorizing the impact of clinical workflow, behavioral, and environmental factors on physiological sensors. *MedStar Health Research Institute Research Symposium*. April 30, 2018, Bethesda, MD (Poster Presentation).
4. Will, A., Kim, T., Fong, A., Kellogg, K. Investigating Factors Related to Emergency Medicine Physician Stress. MHRI 2018 MedStar Health *Research Symposium*. April 30, 2018, Bethesda, MD (Poster Presentation).

5. Kim, T., Ratwani, R., Fairbanks, R.J., Kellogg, K. Identifying environmental challenges and stressors for emergency physicians using a semi-structured survey instrument. *HFES 2019 International Symposium on Human Factors and Ergonomics in Health Care*. March 24, 2019, Chicago, IL (Poster Presentation). Accepted.

Key Words:

Stress, heart rate, emergency medicine