Duty Hours Tracking – Is There an App for That?

Alexis Jorgensen[1], Nastassia M. Savage[2], Xinxin Sun[2] and Gregory Domson[1]

[1]Virginia Commonwealth University Health System, Richmond, Virginia.
[2]Virginia Commonwealth University, Richmond, Virginia.

ABSTRACT

BACKGROUND: To monitor duty hour compliance residency programs have used self-report methods which can be skewed by recall bias and data falsification. The purpose of this study was to compare the accuracy of and resident attitudes towards two duty hours tracking tools within our Orthopedic residency. We compared our institution's current self-report method of duty hours tracking via New Innovations (NI) with an automated method utilizing Hours Tracker (HT), a smartphone application which automatically logs work hours via GPS coordinates. The primary outcome measures were number of duty hour violations and survey results on resident perceptions.

METHODS: The participants were 22 residents of our 25 resident Orthopedic program. Over four weeks, residents tracked duty hours through the standard, selfreport method (NI) and simultaneously through the automated app (HT). Residents also completed an anonymous survey at the end of the study related to perceptions of the methods.

RESULTS: There was no significant difference in overall number of violations between NI and HT. HT detected more violations of the 8 hours off requirement (12 vs. 5, $p = 0.03$). Survey data revealed residents found HT significantly easier to use ($p = .004$) and less burdensome ($p < .001$) but in greater violation of privacy ($p = .001$). Residents reported they were more likely to falsify their hours when using NI ($p = .002$) and that the results of NI would be more likely used against them ($p = .042$). When analyzing by training year, junior residents indicated HT was overall easier to use than senior residents ($p = .048$).

CONCLUSIONS: Our study showed NI and HT are at least equivalent in accuracy with the app being overall better received, particularly by junior level residents. Until we begin accurately tracking duty hours and engaging residents with an easy to use, well-received interface to which report hours, effective developmental program changes will be difficult to achieve. An app-based approach is a starting point for re-thinking duty hours tracking within this digital age.

KEYWORDS: MeSH terms: internship and residency, education, medical, quality improvement

Introduction

Duty hour regulations were initially implemented by the Accreditation Council for Graduate Medical Education (ACGME) in 2003 with the goal of making fundamental changes in resident education and well-being. Due to ongoing concerns of resident fatigue and burnout, the regulations were revised in 2011 to what they are today. Restrictions limit residents' monthly work hours and in-house call as well as require minimum amounts of time off after clinical and educational duties. Surgery residents, especially those training in orthopedics, tend to work more hours per week, get fewer hours of sleep, and have higher rates of burn out in comparison to non-surgical residents.[1]

Enforcement of work hours restrictions is in part attained by requiring program compliance for accreditation. To monitor compliance, residency programs have utilized various methods of duty hours tracking. Most methods use some form of self-reporting whether it be electronic or manual. The accuracy of duty hour self-reporting has been called into question for various reasons including recall bias and data falsification. The true effect of recall bias is unclear. Some studies have shown no significant difference in number of violations between self-report and automated methods.[2–4] Other studies have demonstrated increased violations, upwards of 219%, when utilizing an automated tracking method in comparison to self-report.[5,6] Resident falsification of duty hours also complicates accurate monitoring. Rates of falsification of self-report hours range from 25% to 39%.[5,7,8] Residents who falsify duty hours most commonly do so in fear of personal or departmental repercussions.[8] It is clear we have yet to identify an accurate and well-received method for monitoring duty hours.

As the digital revolution continues to shape the world today, smartphone use particularly among younger generations has become pervasive. A 2016 Nielsen company survey found 96–97% of people aged 25–44 owned smartphones. These devices are an integral part of residents' daily personal and professional lives. Smartphones are now commonly used as a virtual pager, a means to access the medical record, or as a resource for medical information via various applications. The ubiquity of smartphone use creates a unique opportunity for duty hours monitoring not yet explored.

Within our orthopedic surgery residency, duty hours are self-reported through a residency management program called New Innovations (NI). At the time of this study, the GME office was mandating multiple four-week periods of duty hours self-reporting through NI. If a resident logs work hours that violate the ACGME restrictions, an error message is generated within NI. The resident must then provide a justification for the violation. Anonymous yearly program surveys completed by our residents showed a discrepancy between perceived frequency of violation of duty hours and the NI violations counts. This led us to question the accuracy of our own institution's self-report method.

We hypothesized an automated method of duty tracking hours using a smartphone application called "Hours Tracker" (HT) would: 1. Provide more accurate duty hour monitoring 2. Be better received by residents compared to NI. The purpose of this study was to compare the accuracy and resident perceptions of two duty hours tracking tools, NI and HT, over a four-week period within our Orthopedic residency program. The primary outcome measures were number of duty hour violations and survey results on resident perceptions.

Materials and Methods

Participants

The participants in this longitudinal study were the residents of a five-year Orthopedic Surgery program in an academic medical center in the southeastern United States. Twenty-two of the 25 residents (the author was excluded) volunteered to participate in this quality improvement study.

ACGME duty hours restrictions

All violations for this study are based on the five work hour restrictions which are outlined in the Clinical Experience and Education section of the ACGME Common Residency Program Requirements. (https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRResidency2020.pdf).

Hours tracking tools

The study period coincided with a GME mandated four-week duty hours collection period from November 18, 2019 to December 15, 2019. The residents were required by the GME to track duty hours in NI during this period. Twenty-two residents agreed to participate in the study and were asked to concurrently track hours with the HT app. Of the twenty-two residents who used both NI and HT during the study period, two residents had to be excluded due to insufficient (<2 weeks) data. One resident reported phone connectivity issues and the other resident was rotating on a non-clinical block which allowed for significant time out of the hospital.

New innovations hours logging

NI is a residency management program utilized by our GME office which allows residents to self-report hours. Residents must log one-hour blocks of time within the application. Each hour logged must also be categorized as either in-house call, clinical duties, or education. Should the resident log hours which violate one of the ACGME duty hours restrictions, an error message appears stating one or more of the logs are "in violation." The resident is provided an explanation of which rule was violated and then required to select a cause for the violation and comment on details of the violation.

Hours tracker application

HT is a free iOs application available for smart phones. It allows for automated logging of work hours by enabling the user to set "work sites" through GPS coordinates. With location services enabled, the application automatically clocks the user in and out of work whenever the phone is detected entering or leaving pre-designated work sites. The application can generate an exportable data spreadsheet of raw hours worked at each work site.

Prior to the start of the collection period, the residents participated in a tutorial on the application. Residents were provided instructions on downloading the app and creating work sites. The residents were instructed to enter the three most common work locations for the upcoming four-week collection period. For example, a resident may have entered our institution's main hospital, the Veterans Affairs Hospital, and a commonly frequented outpatient clinic for his/her given clinical rotation. GPS coordinates were provided to the residents for all possible work sites as to standardize location parameters. At the end of the four-week duty hours collection period, the residents were asked to export the raw work hours to the author.

Survey collection

All residents were invited to complete an anonymous electronic survey at the end of the four weeks related to perceptions towards the monitoring methods. Of the 22 eligible residents, 4 PGY-1, 3 PGY-2, 5 PGY-3, 4 PGY-4, and 4 PGY-5's responded to the survey, an overall response rate of 91% (n = 20). To ensure confidentiality of responses, the first two training years were combined into a junior level residency group and years 3–5 were combined into a senior level residency group for analyses purposes.

IRB approval

The Institutional Review Board was consulted. The study was deemed quality improvement and exempt from approval.

Measures

NI generates individual resident reports with raw hours worked, hours worked over the four-week period, as well as number and type of duty hours violations incurred during the collection period. HT provides raw hours worked at each work site over the four-week period. The HT raw hours data was then used to generate identical measures as NI output including hours worked over the four week period, number of violations per residents, and type of violation incurred. Time stamps between 0 and 7 minutes on a day without any other logged work hours were assumed to be inadvertent "clock ins" caused by the resident being near a work site or passing by on foot or by car. This time was not counted towards work hours or violations.

The survey items developed by the authors can be found in Table 1. To minimize respondent fatigue, one item was used to assess both NI and HT. Because each question only used one item, reliability of the items was unable to be determined.

Analyses

A Wilcoxon signed rank test was applied to the tracking tool data to determine whether the total number of violations differed between the two systems. Due to data sparsity, the count data of individual violation categories were dichotomized to either 0 or 1, with 0 being no violation and 1 indicating at least one violation. McNemar's test was used on the dichotomized data to test the difference in the number of residents who had at least one violation across the five violation categories.

To compare violation counts among the residents based on training level, two groups were created – junior residents (Post Graduate Year (PGY) 1 and 2) and senior residents (PGY-3 through 5). A Wilcoxon exact rank sum test was utilized. The data was again dichotomized, and a Fisher's exact test

was used to test the difference in number of residents in junior verses senior levels with at least one violation across the five violation categories.

The statistical software used for data management and plotting was R 3.6.1. The software for data analysis were R 3.6.1, SPSS 26.0.0.1, and SAS 9.4.

To analyze the survey data, a paired samples t-test was used to test for differences between the resident perceptions of each tool. To determine whether these differences were impacted by training year, a repeated measures ANOVA was used (Table 2).

Results

Hours tracking results

Violation type and counts within each tracking method are found in Table 1. The overall number of violations did not differ significantly between NI and HT over the study period for all residents. However, there was a trend for HT to detect more violations than NI (20 vs. 12, p = 0.1172) (Figure 1). In analysis of the dichotomized data, HT recorded more violations of the "< 8 hours off between clinical/educational activities" rule than NI (12 vs. 5, $p = 0.0313$) (Figure 2). There were no differences in number of residents who had at least one violation between HT and NI for all other violation categories. When analyzing by training level, there was no significant difference in overall number of violations between HT and NI when comparing junior and senior level residents ($p = 0.1001$; $p = 0.2105$) (Figure 3) or across all violation types (Table 3).

Survey results

There was a significantly higher perception that the results from NI would be used against the individual residents than those from the HT app ($p = .042$). However, there was a significantly

Table 1. Paired samples T-test and descriptives by tracking tool – mean and standard deviations for each survey item are provided for each tracking tool along with results of the paired samples t-test.

ITEMS*	NEW INNOVATIONS MEAN (SD)	HOURS TRACKER MEAN (SD)	PAIRED SAMPLES T-TEST
I believe that the results of duty hours tracking (eg, violations) will be used against me.	**2.74 (1.15)**	2.53 (1.12)	**t(18) = 2.191, p = .042**
I believe that the results of the duty hours tracking (eg, violations) will be used against my program.	3.53 (1.12)	3.32 (1.29)	t(18) = 1.166, p = .259
I feel my privacy is violated by using the _____ duty hours tracking program.	2.37 (1.12)	**3.21 (1.36)**	**t(18) = −3.831, p = .001**
It is easy to use _____ to track my hours.	2.95 (1.03)	**4.16 (1.12)**	**t(18) = −3.260, p = .004**
Tracking my hours in _____ is burdensome.	**3.63 (0.90)**	2.00 (1.11)	**t(18) = 4.429, p < .001**
I like using _____ to track hours.	2.68 (0.89)	3.32 (1.16)	t(18) = −1.528, p = .144
The data in _____ accurately portrays the hours I spent working.	3.42 (0.84)	3.53 (1.27)	t(18) = −0.309, p = .761
I falsify my duty hours in _____.	**2.26 (0.99)**	1.58 (0.77)	**t(18) = 3.367, p = .002**

*Blanks in the items indicate tracker tool name.

Table 2. Data summary of violation types and counts - violations summarized as frequencies and column percentages.

	HT			NI		
	OVERALL (N = 20)	JUNIOR LEVEL (N = 10)	SENIOR LEVEL (N = 10)	OVERALL (N = 20)	JUNIOR LEVEL (N = 10)	SENIOR LEVEL (N = 10)
Violation Type						
> 80 WORK HOURS / 4 WEEKS	1 (5%)	1 (6.25%)	0 (0%)	1 (8.3%)	1 (8.3%)	0 (0%)
< 8 hours off	12 (60%)	9 (56.25%)	3 (75%)	5 (41.7%)	5 (41.7%)	0 (0%)
< 14 hours off	4 (20%)	4 (25%)	0 (0%)	2 (16.7%)	2 (16.7%)	0 (0%)
< 1 day off in 7 days	1 (5%)	1 (6.25%)	0 (0%)	1 (8.3%)	1 (8.3%)	0 (0%)
> 28 hours continuous work	2 (10%)	1 (6.25%)	1 (25%)	3 (25%)	3 (25%)	0 (0%)
Total violations	20	16	4	12	12	0
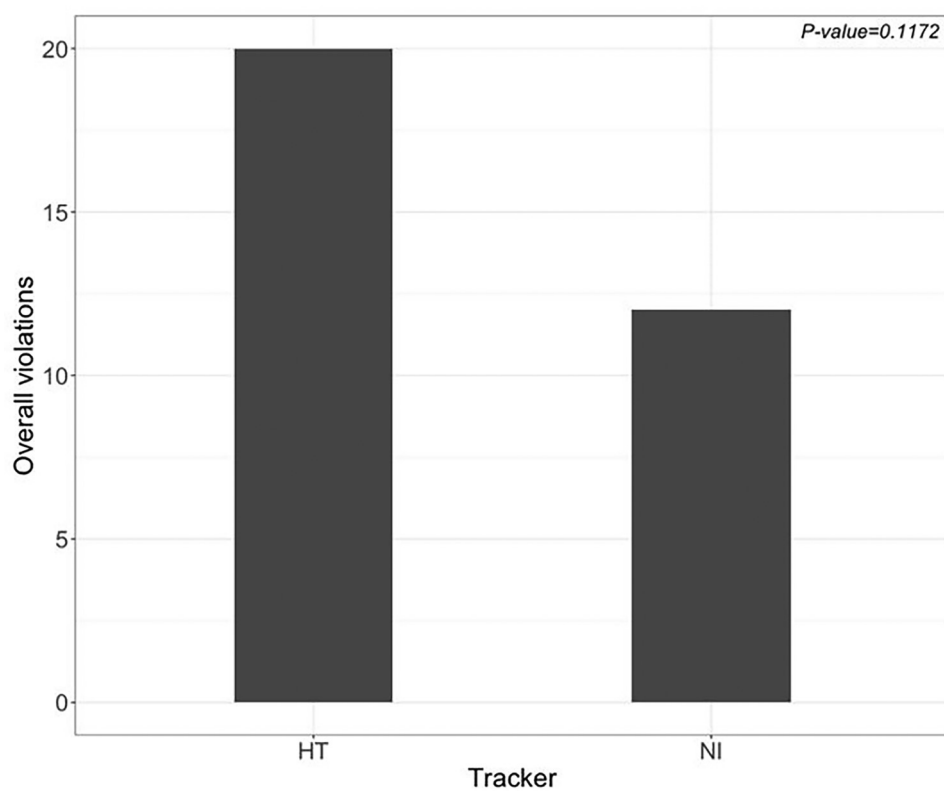
Figure 1. Overall violation counts – HT versus NI (20 vs 12, p = 0.1172).

higher perception that the HT application violated residents' privacy compared to NI ($p = .001$). The HT tool was also significantly easier to use ($p = .004$) and less burdensome than NI ($p < .001$). The residents were significantly more likely to falsify their duty hours reports when using the NI method ($p = .002$) (Table 1).

To determine whether training year impacted the survey results, a repeated measures ANOVA was run with each of the item pairings. Although most of the relationships were not significantly impacted by training year, junior residents perceived HT as significantly less burdensome than senior

residents ($p = .048$) (Figure 4). There were several notable trends. First, junior residents indicated HT to be significantly less likely to be used against them than NI ($p = .082$) (Figure 5). Additionally, the junior residents found the HT tool to be substantially easier to use than the senior level residents ($p = .052$) (Figure 6).

Discussion
Our aim was to explore the accuracy of and resident attitudes towards an automated, smartphone app hours tracking method in comparison to our institution's current self-report
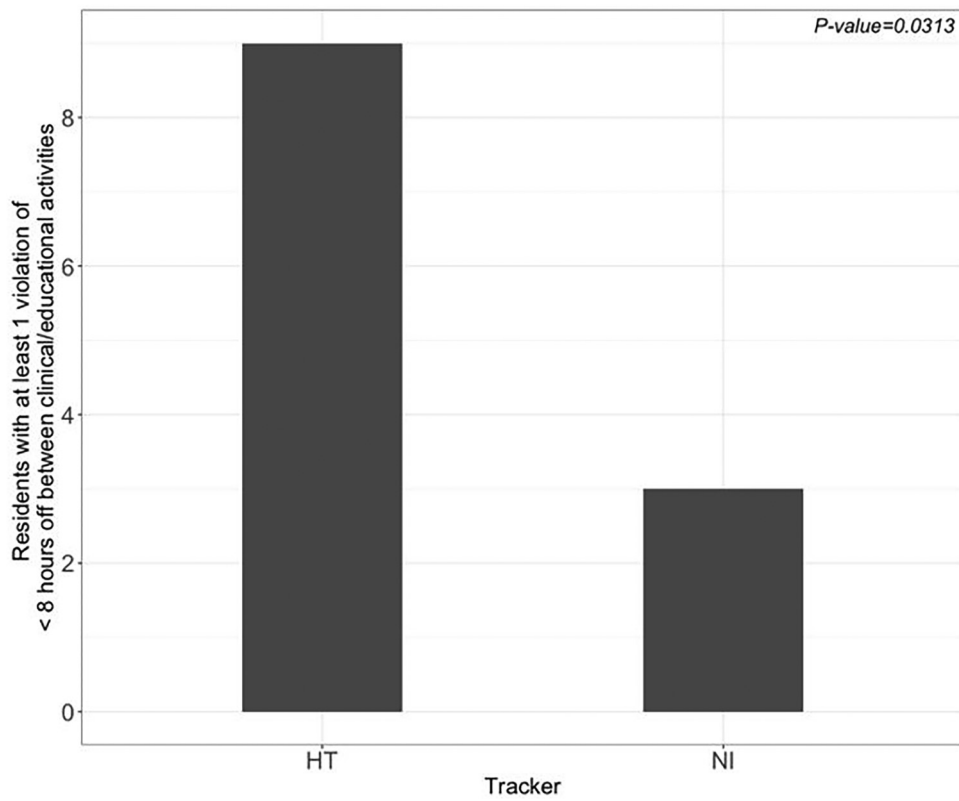
Figure 2. Number of residents with at least one violation of <8 hours restriction – HT versus NI (12 vs 5, *p* = 0.0313).
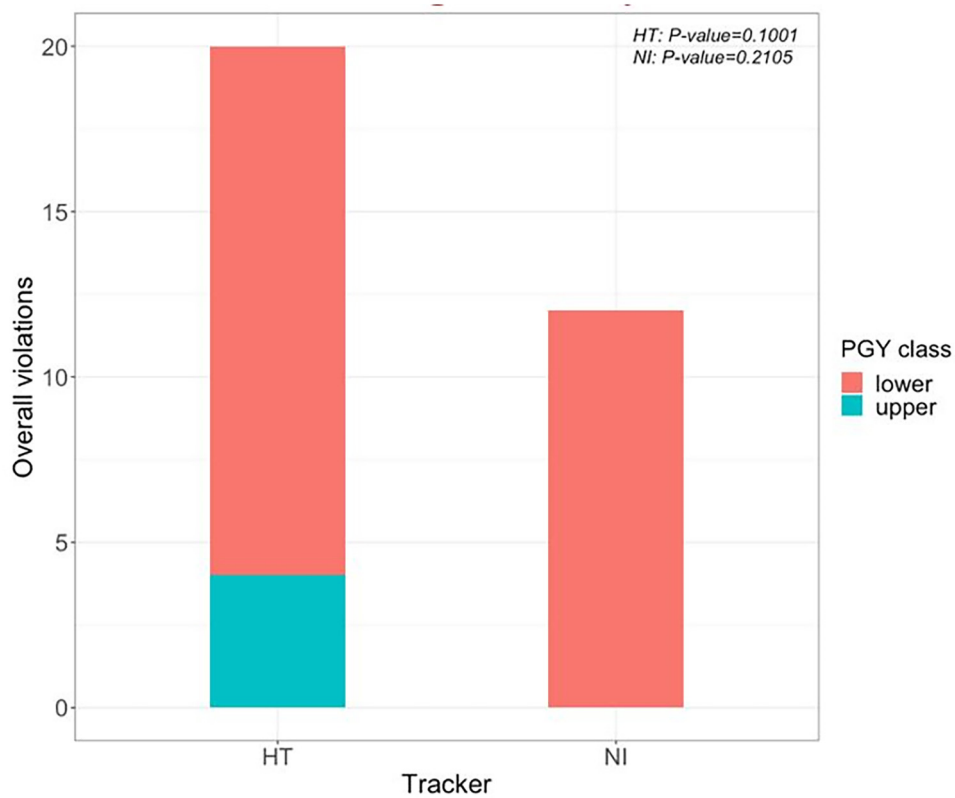

Figure 3. Overall violation counts between junior and senior level residents for HT and NI (HT 16 vs 4, p = 0.1001; NI 12 vs 0, p = 0.2105).

program. Our analyses showed HT to be at least equivalent in accuracy to NI in tracking overall violation counts. HT may be superior to self-reporting in some regards as it detected more violations of the "short break rule" (at least 8 hours off between educational/clinical duties) similar to Buum et al's findings.

In the surveys residents indicated they were more likely to falsify hours within the self-report system. This may account for the increased violations noted in the HT data for short breaks. It is difficult to falsify large chunks of time such as excluding a 24-hour call shift or trim down ten hours of work per week to comply with the 80-hour restriction. However,

residents likely find it easy to slightly adjust their work start and end times day to day and after in-house call to fit the 8-hour and 14-hour free time minimums, respectively. By falsifying these small periods of time, residents avoid a violation error in NI, the burden of justifying the violation, and any perceived professional repercussions.

Overall, HT was better received by residents who indicated the app was easier to use, less burdensome, and less likely to be used against them in comparison to the standard self-report system. Interestingly, training year impacted perceptions. Junior level residents tended to better receive the app to a greater degree than their senior level colleagues. Differences in attitudes towards duty hour restrictions between junior and senior level residents has been well documented. [9] It follows there may be differential perceptions towards tracking methods. Traditionally, junior residents are younger and likely to be more apt with smartphone use. Additionally, they may gain greater satisfaction from use of a technologically current means of tracking duty hours than their senior colleagues.

The duty hour monitoring method selected by an institution should achieve multiple goals including but not limited to minimization of recall bias, ease of use, and universal resident participation. [10] From our survey data, an app-based approach to duty hours monitoring may achieve many of these aims. Yet, there is still room for improvement. Residents indicated
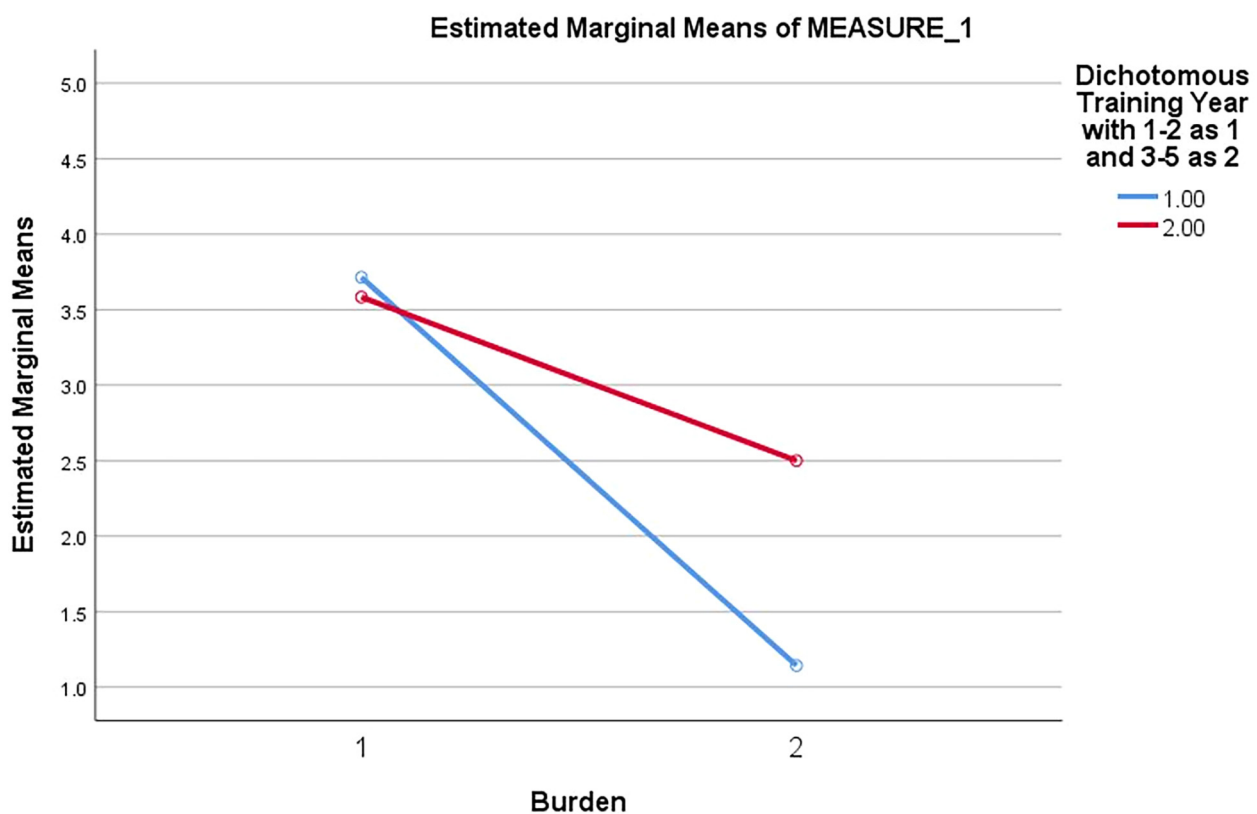
Table 3. Fisher's exact test results for individual violation categories for HT and NI between junior and senior levels.

CATEGORIES	HT	NI
	P-VALUE	P-VALUE
> 80 work hours / 4 weeks	1	1
< 8 hours off between clinical/educational activities	0.3698	0.2105
< 14 hours off between in-house calls	0.2105	1
< 1 day off in 7 days / 4 weeks	1	1
> 28 hours continuous work	1	0.4737

Figure 4. Perceived burden of tracking tool by training year (junior vs senior, p = 0.48). X-axis – tracking tool: 1 – NI, 2 – HT. Y-axis – marginal means of burden. Lines – training year: blue – junior; red – senior.
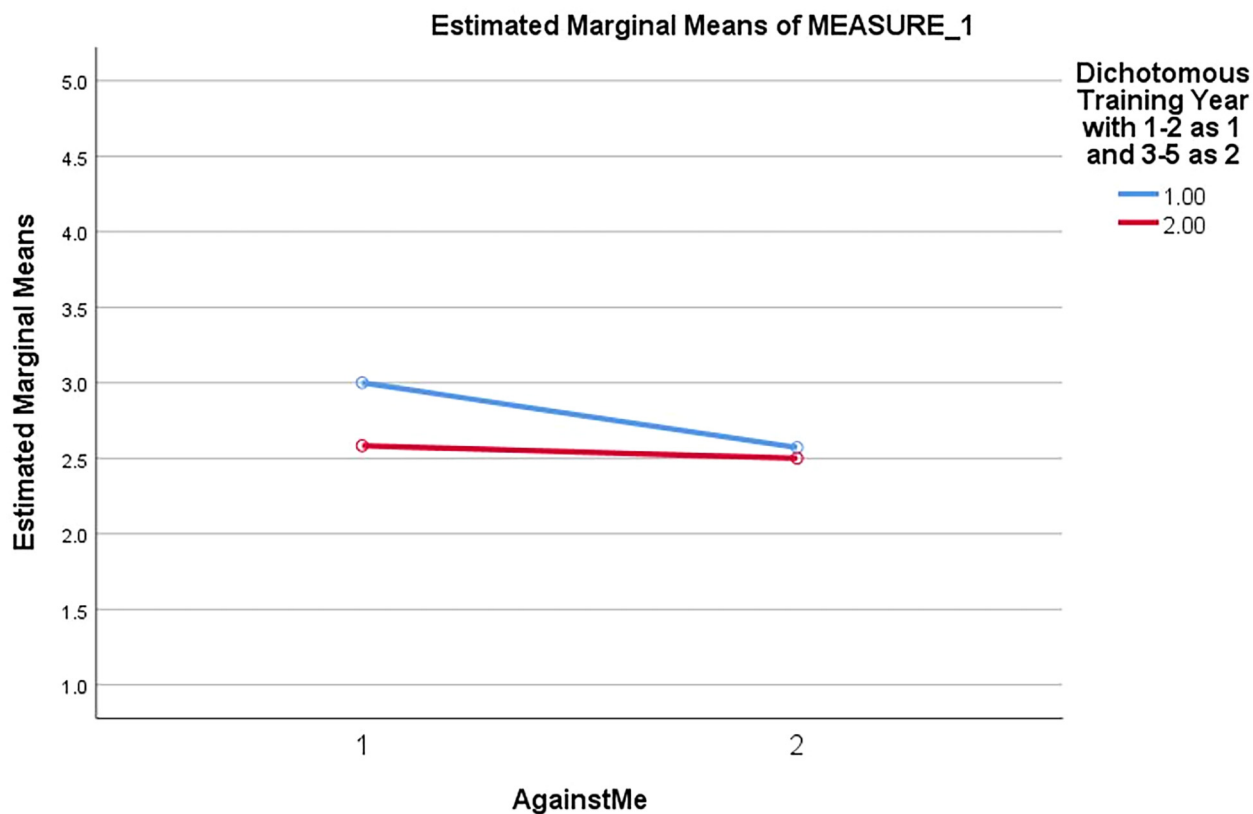
Figure 5. Perceived likelihood results would be used against resident, analyzed by training year (junior vs senior, $p = .082$). X-axis – tracking tool: 1 – NI, 2 – HT, Y-axis – marginal means of likelihood to be used against resident, Lines – training year: blue – junior; red – senior.
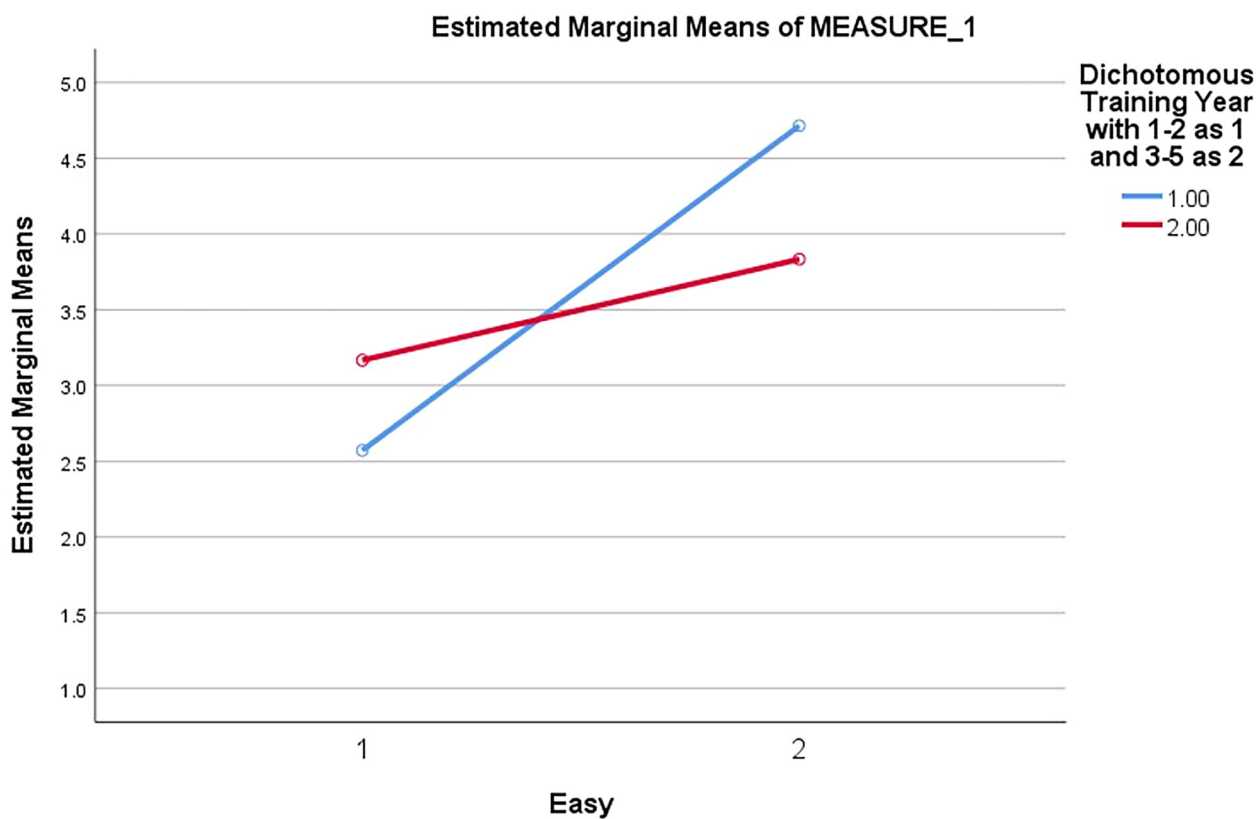

Figure 6. Perceived ease of tracking tool by training year (junior vs senior, $p = .052$). X-axis – tracking tool: 1 – NI, 2 – HT, Y-axis – marginal means of ease of use, Lines – training year: blue –junior; red – senior.

concerns with privacy violation as HT uses location-based services to automatically clock the user in and out. We must prioritize resident perceptions and use this feedback to continue improving upon our duty hours tracking methods. Until we begin accurately tracking duty hours and engaging residents with an easy to use, well-received interface to which report hours, effective developmental program changes will be difficult to achieve.

Our study had several limitations. We had a small sample size (n = 20) over a short period of time (4 weeks) which affected the power of our results. To compare counts of violations in light of the small sample size, the residents were grouped into junior and senior levels and the data was dichotomized. Additionally, the HT application data was not complete for all residents included in the study. The minimum for inclusion was at least two weeks of data. Due to occasional phone connectivity issues and/or delays in setting up the application, not all residents had hours tracked within the application for true workdays. There were also occasional lapses in time at the same work site, up to 2 hours and 12 minutes at the longest, which were assumed to be caused by connectivity issues and combined into continued hours worked. In contrast, the NI data was complete for all days worked as required by our program and GME office. This certainly may have affected the violation counts. Finally, survey data inherently has limitations which we attempted to minimize with paired samples tests to detect differences in perceptions within individuals.

Our findings are specific to our program and institution and cannot be generalized. However, this exploratory study is an important starting point for re-thinking duty hours tracking methods within this digital age while prioritizing resident perceptions as younger generations of orthopedic surgeons enter training. We plan to repeat our study for a longer period to increase the reliability of our findings. Future studies might look at optimization of an app-based approach by mitigating privacy issues. Future studies might also include different orthopedic programs and other residency specialties to help attain a more systematic weighing of pros and cons of various available automatic and self-report tracking tools.

Conclusion

A four-week study directly comparing our institution's self-report duty hours tracking system, NI, with an automated app-based approach demonstrated the app to be at least as accurate as NI in tracking violations. Based on survey findings, resident perceptions towards the app were more positive. This study reveals an app-based approach to duty hours monitoring may be an acceptable alternative to accurately tracking duty hours violations and overall be better received than self-report methods.

Ethical Approval
Not applicable, because this article does not contain any studies with human or animal subjects.

Informed Consent
Not applicable, because this article does not contain any studies with human or animal subjects.

Trial Registration
Not applicable, because this article does not contain any clinical trials.

REFERENCES
1. Mendelsohn D, Despot I, Gooderham P, Singhal A, Redekop G, Toyota B. Impact of work hours and sleep on well-being and burnout for physicians-in-training: the resident activity tracker evaluation study. *Med Educ*. March 2019;53(3):306-315.
2. Todd SR, Fahey BN, Paukert JL, Mersinger D, Johnson M, Bass B. How accurate are self reported resident duty hours? *J Surg Educ*. 2010;67(2):103-104.
3. Shine D, Pearlman E, Watkins B. Measuring resident hours by tacking interactions with the computerized record. *AM J Med*. 2010;123(3):286-290.
4. Chadaga SR, Keniston A, Casey D, Albert RK. Correlation between self-reported resident duty hours and time-stamped parking data. *J Grad Med Educ*. 2012;4(2):254-256.
5. Buum H, Duran-Nelson AM, Menk J, Nixon LJ. Duty-hours monitoring revisited: self-report may not be adequate. *Am J Medicine*. 2013;126(4):362-365.
6. Blitz JB, Rogers AE, Polmear MM, Owings AJ. Duty hour compliance: a survey of U.S. Military medical interns and residents. *Mil Med*. 2017;182(11):e1997-e2004.
7. Zamani N, Fallon SC, Bremer A, Silberfein EJ, Rosengart TK, Scott BG. Using an anonymous, resident-run reporting mechanism to track self-reported duty hours. *Am J Surg*. 2019;218(1):225-229.
8. Byrne JM, Loo LK, Giang DW. Duty hour reporting: conflicting values in professionalism. *J Grad Med Educ*. 2015;7(3):395-400.
9. Mauser NS, Michelson JD, Gissel H, Henderson C, Mauffrey C. Work hour restrictions and orthopaedic resident education: a systematic review. *Int Orthop*. May 2016;40:865-873.
10. Chang LW, Vidyarthi AR, Kohlwes RJ. Baseline duty hours recorded with time-cards: a pre-regulation study of internal medicine residents. *Med Educ*. 2006;40(7):662-667.