

Innovation to Improve GME: Automation Increases Scheduling Efficiency and Improves Schedule Quality for Medical Residents

Elizabeth Perelstein¹; Ariella Rose²; Young-Chae Hong, MSE²;
Amy Cohn, PhD²; Micah T. Long, MD³

¹ University of Michigan Medical School, Ann Arbor, MI, USA.

² Department of Industrial & Operations Engineering, University of Michigan College of Engineering, Ann Arbor, MI, USA.

³ Department of Pediatrics and Communicable Diseases, University of Michigan's C.S Mott Children's Hospital, Ann Arbor, MI, USA.

Elizabeth Perelstein is a third-year medical student at the University of Michigan Medical School, Ann Arbor MI.

Ariella Rose is a recent graduate of the Department of Industrial and Operations Engineering, University of Michigan College of Engineering, Ann Arbor MI.

Young-Chae Hong is a PhD candidate at the Department of Industrial and Operations Engineering, University of Michigan College of Engineering, Ann Arbor MI.

Dr. Amy Cohn is the Associate Director of the Center for Healthcare Engineering and Patient Safety, Ann Arbor MI, and an Arthur F. Thurnau Associate Professor of the Department of Industrial and Operations Engineering, University of Michigan College of Engineering, Ann Arbor MI.

Dr. Micah T. Long is a recent graduate of the Pediatrics Residency Program, Department of Pediatrics and Communicable Diseases, University of Michigan's C.S. Mott Children's Hospital, Ann Arbor MI.

Funding/Support: This work was supported by grants from The University of Michigan Center for Healthcare Engineering and Patient Safety, The Seth Bonder Foundation, and The Doctors Company Foundation.

Correspondence should be addressed to Dr. Cohn, Dept. of Industrial & Operations Engineering, 1205 Beal Avenue, Ann Arbor MI 48109. E-mail: amycohn@umich.edu. Tel.734.615.7258. Fax 734.764.3451.

ABSTRACT

Purpose: Medical resident scheduling is a difficult task due to multiple sets of rules, competing educational goals, and ever-evolving GME requirements. Despite this, schedules are typically created by hand, consuming hours of work, producing schedules of varying quality, and yielding negative consequences for resident morale and learning. The objective of this research is to determine whether a computerized scheduling assistant can improve schedule quality and decrease time to schedule completion, thereby saving residency programs time, saving money, and improving morale.

Methods: The “Optimized Residency Scheduling Assistant” (ORSA) was designed by a team from the University of Michigan Department of Industrial and Operations Engineering. It was implemented in the C.S. Mott Children’s Hospital pediatric emergency department in the 2012-2013 academic year. Four metrics of schedule quality were compared between the 2010-2011 and 2012-2013 academic years, including incidence of challenging shift transitions, incidence of shifts following continuity clinics, total shift inequity, and night shift inequity.

Results: All scheduling rules were successfully incorporated. Average schedule creation time fell from 22-28 hours to 4-6 hours per month and three out of four metrics of schedule quality were significantly improved. Challenging shift transitions decreased by an average of 0.54/resident/month (-85.7%, $p < 0.01$); post-clinic shifts decreased by 0.36/resident/month (-66.7%, $p < 0.01$); and the standard deviation of night shifts dropped by 0.05/resident/day (-44.4%, $p < 0.01$).

Conclusions: Automated shift scheduling improves the current manual scheduling process, reducing time spent and improving schedule quality. Embracing such automated tools can benefit residency programs with shift-based scheduling needs.

Innovation to Improve GME: Automation Increases Scheduling Efficiency and Improves Schedule Quality for Medical Residents

INTRODUCTION

Medical residents have unique and complex scheduling needs related to their training. There are hospital- and program-specific rules governing the shift schedule in addition to personal preferences for days off and vacation. Equally importantly, residents must abide by the Accreditation Council for Graduate Medical Education (ACGME) enforced work-hour restrictions, which include guidelines regarding days off, time between shifts, maximum shift duration, and the number of consecutive night shifts. Residents also must work a minimum number of shifts to receive adequate training, take part in varied educational activities and conferences outside of the hospital, and provide medical care to the community at local clinics. With continuing calls to reform GME, this already-complicated picture is likely to grow increasingly complex with time.

Designing a high-quality resident schedule by hand which addresses both ACGME scheduling requirements and resident preferences is difficult, time-consuming, and error-prone. By hand, it is difficult to achieve even a feasible schedule, here defined as one that satisfies all strict requirements. This is particularly so in environments staffed by several residency programs simultaneously: for example, the pediatric emergency department (ED) environment, which is staffed by residents from pediatrics, family practice, and emergency medicine. Each set of residents has unique experience, abilities, educational goals, and out-of-hospital requirements including community clinic responsibilities. The differences between program requirements are

difficult to address by hand and force the scheduler to focus on mere feasibility, decreasing the emphasis on schedule quality, resident education, and patient care.

Yet a poor-quality resident shift schedule can yield negative consequences for both patients and staff. Uneven shift distributions result in poor morale, raising the risk for resident burnout and resulting in an overall sub-optimal educational environment. Poor quality schedules can further negatively impact residents via fluctuations in daily scheduled sleep periods, as dictated by a resident's time between shifts, forcing residents to work against their circadian rhythms. This lowers the magnitude of physiological factors related to wakefulness, including melatonin levels and body core temperature changes throughout the day, ultimately contributing to resident fatigue.[6] Fatigue is a profound problem in medical residency programs, and has been shown to depress fine-motor skills and cognition.[2, 3, 11, 12] Fatigue can also present a danger to residents outside of the hospital. In addition to the negative effects of sleep loss and fatigue on residents' personal lives, residents experiencing excessive fatigue have an increased risk of negative health events, including motor vehicle accidents.[8, 9, 10] Poor scheduling, therefore, places the hospital at risk for both diminished patient satisfaction and adverse health events for staff.

Smarter staff scheduling must be explored. Both residents and the patients they treat would benefit substantially from a computerized tool to automate the scheduling process, addressing all of the system complexity and efficiently generating high-quality schedules. To address these challenges, the pediatric ED of the University of Michigan C.S. Mott Children's Hospital collaborated with the Center for Healthcare Engineering and Patient Safety at the University of

Michigan's College of Engineering. In an attempt to substantially improve both the creation process and the quality of schedules for the pediatric ED residents, a computerized optimization-based scheduling tool was developed: the Optimized Residency Scheduling Assistant (ORSA). The program is a unique addition to the world of medical scheduling tools and may be better equipped to deal with evolving policies in GME. Most current scheduling tools optimize one certain parameter of that schedule, such as shift preference or cost.[4,5,7,13] However, multiple important criteria often exist in medical resident scheduling, and a multi-criteria function is more appropriate. Additionally, the relationships between criteria are difficult for schedulers to quantify and often change from month to month; a mathematically "optimal" solution therefore may not exist.[1] Accordingly, ORSA allows the user to adjust several metrics as he or she sees fit, according to the needs of that specific month.

ORSA ensures that all rules and requirements are met, while enabling the scheduler to easily impose preferences that improve shift equity between residents, provide the hospital with more experienced residents during critical hours, and improve resident sleep patterns both in the hospital and in community clinics.

METHOD

Model creation

The scheduling tool itself was comprised of an integer programming model. Decision variables corresponded to the question of whether or not to assign a particular resident to a particular shift on a particular day. One such variable was defined for each resident per shift per day

combination. Mathematical constraints then enforced scheduling rules and relationships; for example, a constraint enforced the rule that residents must have at least ten hours off between consecutive shifts. The system was implemented and solved using C++ with IBM's CPLEX API v12.1 on a computer with an Intel Xeon 3.20 GHz processor and 8 GB of memory. Run times for each iteration (i.e. the time required by the computerized tool to create a schedule given a set of input files) were measured.

Study design and setting

The data was collected between January and July 2013 in the University of Michigan C.S. Mott Children's Hospital Pediatric Emergency Department.

Records of pediatric emergency department resident schedules were reviewed, dating from July 2010 through June 2011 and from July 2012 through June 2013 (the 2010-2011 and 2012-2013 academic years, respectively). The 2010-2011 year was the most recent year that the schedule was constructed completely by hand and the 2012-2013 year was the first complete year scheduled using ORSA. The intervening 2011-2012 academic year was a year of transition and was therefore omitted from analysis.

During both of the study years the pediatric ED had six required resident shifts: 7:00 AM - 4:00 PM, 9:00 AM - 6:00 PM, 4:00 PM - 1:00 AM, 5:00 PM - 2:00 AM, 8:00 PM - 5:00 AM, and 11:00 PM - 8:00 AM. Two additional optionally-filled shifts, one from 10:00 AM - 7:00 PM and the other from 12:00 PM - 9:00 PM, were filled as frequently as possible.

Residents assigned to the pediatric ED during any given month spent between one and four weeks on service. Resident groups that worked shifts in the ED included pediatrics, combined medicine and pediatrics, emergency medicine, and family practice. Educational requirements outside of the ED differed for each group. Pediatric residents worked at a community clinic one half-day per week on varied days, towards the goal of continuity with their prior established weekly clinic schedule. These generally occurred from 1:00 – 5:00 PM with rare clinics from 9:00 AM – 12:00 PM. Emergency medicine residents were expected to attend weekly educational activities on Wednesdays from 10:00 AM to 2:00 PM. Family practice residents included first-year residents who worked at a community clinic on Wednesdays from 9:00 AM to 5:00 PM, and third-year residents who worked at a community clinic on Mondays and Wednesdays from 9:00 AM to 5:00 PM.

To be feasible, schedules had to allow each resident to attend their outside requirements, described above, and abide by ACGME duty-hour restrictions. These rules dictated that residents have a monthly minimum of four 24-hour periods of time off per week, a maximum of 80-hours worked per week, a minimum of 10-hours between separate clinical shifts or responsibilities, and a maximum of 6-nights worked consecutively.

In the first phase of ORSA-aided scheduling, data was provided to the program including a resident list with clinic days, vacations and home program included. ORSA then automatically generated a feasible schedule. The scheduler then reviewed the schedule to identify undesirable characteristics. In the second phase of scheduling, the scheduler added additional requests, and ORSA attempted to incorporate these additional requirements. It either identified a new,

improved schedule that met the requested specifications or guaranteed with certainty that it was not possible to satisfy all of the requests and still provide a feasible schedule. The second phase typically involved one to two hours of iteratively adding requests and generating new schedules.

This quality improvement project was exempt from IRB oversight.

Study measurements and data

To assess schedule quality, four specific measures were evaluated: total shift disparity, night shift disparity, occurrence of shifts immediately following community clinic responsibilities (“post-clinic shifts”), and occurrence of challenging shift transitions, nicknamed “bad sleep patterns” (BSPs). Total shift disparity and night shift disparity refer to variance in numbers of shifts among residents in any given month. Post-clinic shifts were chosen as a negative quality metric because of the difficulty of preceding a shift with outside clinical requirements. Bad sleep patterns were defined as consecutive shift assignments that yield a negative sleep schedule for residents, and were determined by informally surveying senior residents on challenging shift transitions. These BSPs are listed in Table 1.

Shift Combinations Yielding Bad Sleep Patterns	
1. Day 1: 12 PM – 9 PM,	Day 2: 7 AM – 4 PM
2. Day 1: 4 PM – 1 AM,	Day 2: 12 PM – 9 PM
3. Day 1: 5 PM – 2 AM,	Day 2: 12 PM – 9 PM
4. Day 1: 8 PM – 5 AM,	Day 2: 4 PM – 1 AM
5. Day 1: 8 PM – 5 AM,	Day 2: 5 PM – 2 AM
6. Day 1: 8 PM – 5 AM,	Day 3: 7 AM – 4 PM
7. Day 1: 11 PM – 8 AM,	Day 3: 7 AM – 4 PM
8. Day 1: 11 PM – 8 AM,	Day 3: 9 AM – 6 PM

Table 1: Bad sleep patterns, defined. Consecutive shifts reflecting BSPs are shown.

Data collected included the monthly numbers of residents working in the ED, total shifts per resident, night shifts per resident, post-clinic shifts, and BSPs. Department of Pediatrics chief residents were informally surveyed on the amount of time necessary to create a schedule by hand, both utilizing ORSA and preceding the creation of the tool. Finally, ORSA's computation time to create a schedule was extracted for each month in the 2012-2013 academic year.

Analysis

Extraction of data was fully normalized by month (e.g. such that a resident working in the ED for two weeks of a four-week month was counted as 0.5 residents). Statistical analysis was completed using Minitab® 16 (Minitab Inc., State College, PA, USA).

RESULTS

Informal surveys of former pediatric ED chief residents revealed that a typical schedule took 12 to 16 hours to build by hand, plus another 10 to 12 hours of later corrections. Using ORSA, the schedule took between two to three hours of file-building, plus two to three hours of modification and fine-tuning with the engineering team. There was little-to-no error correction time needed. In sum, the total time to build a schedule was between 22 and 28 hours per month by hand, and four to six hours per month utilizing ORSA. For all months in the 2012-2013 academic year, ORSA's computation time to create a schedule iteration was less than two seconds, with a mean of 0.54 (SEM = 0.15) seconds.

Bad sleep patterns, post-clinic shifts, and measures of shift disparity were calculated for each month in the 2010-2011 and 2012-2013 academic years, then averaged within each year.

Student's *t*-tests compared the year-averaged data between the 2010-2011 and 2012-2013 academic years, and paired *t*-tests compared the month-averaged data between the years (Table 2).

	Academic Year 2010-2011				Academic Year 2012-2013				Difference			
	<i>Total</i>	<i>Mean per resident per mo.</i>	<i>SD</i>	<i>SEM</i>	<i>Total</i>	<i>Mean per resident per mo.</i>	<i>SD</i>	<i>SEM</i>	<i>Mean Δ per resident per mo.</i>	<i>% change</i>	<i>Unpaired p-value</i>	<i>Paired p-value</i>
Bad Sleep Patterns	83	0.63	0.28	0.08	14	0.09	0.31	0.09	-0.54	-85.7%	< 0.01	< 0.01
Shifts After Clinics	72	0.54	0.26	0.08	32	0.18	0.22	0.06	-0.36	-66.7%	< 0.01	< 0.01
Total Shift Variability <i>SD in Shifts / Day</i>	–	0.08	0.02	0.01	–	0.06	0.03	0.01	-0.02	-25.0%	0.49	0.33
Night Shift Variability <i>SD in Night Shifts / Day</i>	–	0.09	0.03	0.01	–	0.04	0.02	0.00	-0.05	-44.4%	< 0.01	< 0.01

Table 2: Bad sleep patterns, post-clinic shifts, and shift disparity. BSPs, post-clinic shifts, and measures of shift disparity were calculated and averaged both per-month and per-year within the 2010-2011 and 2012-2013 academic years. All variables were normalized for number of residents working in the pediatric ED that month.

There was a mean of 13.44 (SD = 2.81, SEM = 0.57) residents working per month in the ED, normalized per month-period, such that a resident who was working in the ED for 15 days in a 30-day month was counted as 0.5 residents.

The number of BSPs decreased significantly from the 2010-2011 to 2012-2013 academic year, from a mean of 0.63 BSPs per resident per month to 0.09 BSPs per resident per month (mean change -0.54, $p < 0.001$, 95% CI: (-0.29, -0.79), percent change -85.7%). Similarly, the number of post-clinic shifts decreased significantly with the average dropping from 0.54 post-clinic shifts per resident per month to 0.18 post-clinic shifts per resident per month (mean change -0.36, $p < 0.001$, 95% CI: (-0.16, -0.57), percent change -66.7%).

Shift disparity was measured as the standard deviation of shifts per resident per calendar day. While there was no significant difference in total shift disparity between years (mean change -0.02, $p = 0.33$, 95% CI (-0.04, 0.00), percent change -25.0%), there was statistically significant reduction in night shift disparity, from 0.09 to 0.04 (mean change -0.05, $p < 0.001$, 95% CI: (-0.03, -0.07), percent change -44.4%).

When the study years were further compared in month-matched paired t-tests, BSPs, post-clinic shifts, and night shift variability continued to be significantly reduced between years, with mean changes of -0.54, -0.36, and -0.05, respectively (all $p < 0.001$). Total shift disparity was not significantly different between the month-matched data (mean change = -0.02, $p = 0.49$).

Results were assessed for dependence on the number of residents working in the ED per month, a number which fluctuates from month to month. Pearson correlation coefficients were calculated for the linear dependence of BSPs, post-clinic shifts, and shift disparity on the number of residents working in the ED (Table 3). Coefficients were calculated for both academic years individually, as well as for the two years grouped together. All values except one were below 0.50; nine out of 12 interactions had coefficients of less than 0.20, another two out of 12 had coefficients less than 0.50, and the final coefficient had a coefficient of 0.53.

	2010-2011	2012-2013	Grouped
Bad Sleep Patterns	0.10	0.17	0.48
Shifts After Clinics	0.01	0.00	0.16
Total Shift Disparity <i>(SD in Shifts / Day)</i>	0.43	0.00	0.09
Night Shift Disparity <i>(SD in Night Shifts / Day)</i>	0.20	0.05	0.53

Table 3: Pearson’s correlation coefficients comparing metrics and resident compliment.

Pearson’s correlation coefficients are displayed for the linear dependence of BSPs, post-clinic shifts, and measures of shift disparity on the number of residents working per month.

DISCUSSION

The advent of the scheduling tool significantly improved several measures of schedule quality. Bad sleep patterns, post-clinic shifts, and night shift disparity all significantly decreased from 2010-2011 to 2012-2013. The change in total shift disparity was not significant, but this is not particularly surprising; the easiest assessment of schedule quality to check by hand is a count of

each resident's shifts per month, and this is where chief residents often invest time in improving handmade schedules.

These improvements in schedule quality came while *decreasing* the time required for schedule production. Compared to making them by hand, ORSA reduced creation time by over 20 hours per monthly schedule. This was partly attributable to the accuracy of automated schedules.

While handmade schedules typically require hours of later corrections, ORSA schedules required little to no correction.

Pearson correlation coefficients revealed that these significant changes in schedule quality were not due to the monthly resident compliment. The coefficients for the interaction of the metrics and the number of residents per month are remarkably low, explaining less than 20% of most of the measured outcomes. In short, improved outcomes cannot be attributed to increased resident numbers.

The benefits of using computer assistance in physician scheduling are numerous. Using an automated tool like ORSA is financially beneficial to the hospital. Typically the chief resident is tasked with making monthly schedules in the pediatric ED. The chief resident is the most expensive resident to the ED, and their time is the most valuable. Using an automated scheduling tool frees roughly 20 hours per month of the chief resident's time, allowing him or her to spend that time on other duties. This would invaluablely extend the utility of automated tools to faculty scheduling models, where the cost of the faculty scheduler's time exponentially increases from that of residents' time.

An additional benefit of using an automated scheduling tool comes from the impact of resident fatigue on morale and patient care. Residents who have poor sleep habits display cognitive declines in a variety of areas, and are more likely to make mistakes.[2, 3, 6, 9] They are also at increased risk for burnout, especially if their schedule features uneven shift distributions perceived to be unfair. By improving upon these factors, automated scheduling tools like ORSA have the potential to reduce the negative effects of poor sleep on patient care. Additionally, ORSA benefits the residents themselves. Poor sleep patterns and unequal shift distributions can not only lead to resident dissatisfaction, but also to negative health consequences, such as an increased incidence of motor vehicle accidents.[8, 10] The scheduler addresses these variables specifically and therefore cuts down on factors leading to resident fatigue and danger.

Finally, automation with a tool like ORSA allows adaption to each month's different resident complement, scheduling idiosyncrasies, and the optimization preferences of the scheduler. Such flexibility leads to increased efficiency, as the scheduler is not forced to start from scratch when these variables change. Currently, this flexibility is relatively unique in the world of scheduling tools. Most tools are focused on nursing, a field with very different scheduling constraints compared to residency.[5, 7, 13] (For example, nurses typically do not have weekly outside clinic requirements, and the rules outlining legal nursing shift schedules are different than those for residents.) Few tools are sufficiently customizable for evolving GME rules and requirements. A good scheduling tool should be able to incorporate ad-hoc adjustments, prioritizing different metrics of schedule quality according to the needs of that month. Most of what is currently available optimizes a certain parameter of a monthly schedule, such as shift preference or cost. By not optimizing one pre-determined trait, ORSA allows a scheduler to

choose what parameter they want to optimize each month. In short, the use of this automated scheduling tool easily addresses changing scheduling requirements and preferences from month-to-month.

Since measures of resident satisfaction, fine-motor skills, educational value, or patient care were not directly assessed, the scope of this paper is limited to the described schedule quality metrics and we cannot definitively state that our schedule improves patient care or resident morale.

These topics may be a focus of future research. However, it follows logically that a schedule that considers the human factors of shift work may yield improvement in those categories.

Another limitation comes from our sample size, which was limited to two academic years, poorly impacting our statistical power. Finally, bad sleep patterns were defined based on experience and resident feedback; there is no evidence specifically associating these shift transitions with poor patient care or educational outcomes.

In conclusion, automated shift scheduling tools like ORSA improve the current manual scheduling process. Automated schedules are made faster and are of higher quality than schedules made by hand, and they are able to efficiently incorporate changing scheduler preferences from month to month. It is our hope that improving schedule quality in this way can improve patient safety, resident education, and morale. Residency programs should strongly consider adopting such tools to meet the challenging and ever-changing demands of resident shift scheduling.

Acknowledgments: The authors would like to thank Dr. Brian Jordan, the University of Michigan Center for Healthcare Engineering and Patient Safety, The Seth Bonder Foundation, and The Doctors Company Foundation, without whom this work would not have been possible.

Funding/Support: This work was supported by grants from The University of Michigan Center for Healthcare Engineering and Patient Safety, The Seth Bonder Foundation, and The Doctors Company Foundation.

Other Disclosures: Dr. Long, Ms. Perelstein, Ms. Rose, and Mr. Hong have nothing to disclose. Dr. Cohn reports grants from Center for Healthcare Engineering and Patient Safety, University of Michigan, grants from The Seth Bonder Foundation, grants from The Doctors Company Foundation, during the conduct of the study, and personal fees from Analytics Corporation, outside the submitted work.

Ethical Approval: Not applicable.

Disclaimer: None.

Previous Presentations: Portions of this work were previously presented at the Industrial & Systems Engineering Research Conference (6/1/14, Montreal, Canada), the Institute for Operations Research and the Management Sciences (INFORMS) Conference (10/9/14, Minneapolis, MN), and the INFORMS Healthcare Conference (5/25/13, Chicago, IL).

REFERENCES

1. Cohn A, Root S, Kymissis C, et al. Scheduling Medical Residents at Boston University School of Medicine. *Interfaces* 2009;39:186-195.
2. Williamson A, Feyer A. Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. *Occup Environ Med* 2000;57:649-655.
3. Dawson D, Reid K. Fatigue, alcohol, and performance impairment. *Nature* 1997;388:235-237.
4. Sherali H, Rmahi M, Saifee Q. Hospital resident scheduling problem. *Production Planning and Control* 2002;13:220-233.
5. Bard J, Purnomo H. Short-term nurse scheduling in response to daily fluctuations in supply and demand. *Health Care Manag Sci* 2005;8:3315-324.
6. Joffe, M. Emergency Department Provider Fatigue and Shift Concerns. *Clin Pediatr Emerg Med* 2006;7:248-254.
7. Azaiez M, Al Sharif S. A 0-1 goal programming model for nurse scheduling. *Comput Oper Res* 2005;32:491-507.
8. Steele M, Ma O, Watson W, Thomas Jr. H, Muelleman R. The Occupational Risk of Motor Vehicle Collisions for Emergency Medicine Residents. *Acad Emerg Med* 1999;6:1050-1053.
9. Papp, K, et al. The Effects of Sleep Loss and Fatigue on Resident-Physicians: A Multi-Institutional, Mixed-Method Study. *Acad Med* 2004;79:394-406.

10. Philip P, Sagaspe P, Moore N, et al. Fatigue, sleep restriction and driving performance. *Accid Anal Prev* 2005;37:473-478.
11. Smith-Coggins R, Rosekind M, Buccino K, et al. Rotating shiftwork schedules: can we enhance physician adaptation to night shifts? *Acad Emerg Med* 1997;4:951-961.
12. Smith-Coggins R, Rosekind M, Hurd S, et al. Relationship of day versus night sleep to physician performance and mood. *Ann Emerg Med* 1994;24:928-934.
13. Aickelin U, Dowsland K. An indirect genetic algorithm for a nurse-scheduling problem. *Comput Oper Res* 2004;31:761-778.