The effect of call on neurosurgery residents’ skills: implications for policy regarding resident call periods

Clinical article

ARUNA Ganju, M.D.,1 KANAV KAHOL, Ph.D.,2,3 PETER Lee, M.D.,3 NARINA Simonian, B.S.,1 STEVEN J. Quinn, B.A.,1 JOHN J. Ferrara, M.D.,4 and H. Hunt Batjer, M.D.1

1Department of Neurological Surgery, Northwestern University Feinberg School of Medicine, Chicago, Illinois; 2Simulation and Education Training Center, Banner Good Samaritan Medical Center, Phoenix; and 3Department of Biomedical Informatics, Center for Cognitive Ubiquitous Computing, School of Computing and Informatics, 4Arizona State University, Tempe, Arizona

Object. Although fatigue and its effects on surgical proficiency have been an actively researched area, previous studies have not examined the effect of fatigue on neurosurgery residents specifically. This study aims to quantify the effect of fatigue on the psychomotor and cognitive skills of neurosurgery residents.

Methods. Seven neurosurgery residents performed a minimum of 3 and a maximum of 4 sessions of 6 surgical exercises precall and postcall. The simulation exercises were designed to measure a surgeon’s cognitive abilities, such as memory and attention, while performing simulated surgical tasks and exercises that have been previously validated in several studies, including studies measuring the impact of fatigue on general surgery residents. Each exercise measured tool-movement smoothness, time elapsed, and cognitive errors. The change in surgical skills in precall and postcall conditions was assessed by means of an ANOVA, with p < 0.05 considered statistically significant.

Results. The neurosurgery residents did not show a statistically significant difference in their surgical skills between the pre- and postcall states (p < 0.3, p < 0.4, and p < 0.2 for movement smoothness, time elapsed, and cognitive errors, respectively). The mean decrement for all residents in the postcall condition was 13.1%.

Conclusions. Postcall fatigue is associated with a marginal decrease in proficiency during simulated surgery in neurosurgery residents. In a similar study, general surgery residents showed a statistically significant decrement of 27.3% in the postcall condition. The impact of fatigue on different specialties should be further investigated prior to implementation of a national physician work-hour policy.

Key Words • Accreditation Council for Graduate Medical Education • work-hour restriction • neurosurgery resident

Graduate medical education was redefined in 2003 with the ACGME’s implementation of work-hour restrictions. These restrictions imposed a “one size fits all” solution to the issues of resident fatigue and patient safety. Under these guidelines, work hours were limited to 80 hours per week, averaged over a period of 4 weeks. In December 2008, the Institute of Medicine released a report calling for further restrictions to be imposed, including mandatory rest time during residents’ duty shifts. In response, in March 2009, the ACGME convened to critically review the impact of the work-hour restrictions on resident education and patient safety after 5 years of implementation. It was recognized that there is a paucity of data regarding the impact of work-hour restrictions on graduate medical training and patient care. In addition, it was acknowledged that there is tremendous variation in residency training across the medical specialties; a work-hour restriction policy applied indiscriminately to all subspecialties is problematic. However, the report also pointed out that there was a lack of comparative data on the variations between specialties.

It is generally recognized that the various medical subspecialties have unique demands; practitioners must have unique abilities and skill sets to practice successfully. Neurological surgery is recognized as being a medical subspecialty in which timely care, delivered at any time of day or night, can often mean the difference between life and death, function and disability. No data exist regarding the effect of fatigue on the psychomotor and cognitive skills of neurosurgery residents. The onus is on medical practitioners to provide these data. It was our hypothesis that fatigue does not affect the psychomo-
Effect of call on neurosurgery residents’ skills

tor and cognitive skills of medical practitioners equally. Specifically, we were interested in comparing, in a controlled manner, the impact of fatigue on the psychomotor and cognitive skills of neurosurgery and general surgery residents. Simulation environments provide a safe and effective means of measuring the impact of fatigue on skills. They have been employed in several studies, and it has been shown that performance in simulation correlates highly with intraoperative performance.8

In a meta-analysis of 19 studies conducted by Pilcher and Huffcutt,9 partial sleep deprivation (< 5 hours of sleep in a 24-hour period) was found to cause a significant impact on overall proficiency of residents. This study noted that cognitive skill (as measured through questionnaires, simple cognitive exercises, and so forth) is more deeply affected by fatigue than is psychomotor skill. In separate experiments, Grantcharov et al.10 and Taffinder et al.11 examined the deterioration of laparoscopic skills in the fatigued condition as measured by the ProMIS surgical simulator (Haptica, Ltd.). In both of these studies, fatigued subjects demonstrated impeded accuracy and increased time span for completion of tasks. Neither study evaluated the effect of fatigue on neurosurgery residents; in addition, both of these studies isolated psychomotor proficiency as a measurement variable instead of recognizing the interrelatedness of psychomotor and cognitive proficiency.

Kahol et al.7 addressed this problem in a recent article; they were interested in developing, for surgeons, exercises and tools that require the use of psychomotor and cognitive skills simultaneously. In this study using a specially designed simulator to measure psychomotor and cognitive skills in general surgery residents, the authors developed an experiment that analyzed the effect of fatigue and sleep deprivation on trauma surgery and obstetric residents. Each participant was involved in 4 precall and 4 postcall sessions. Three exercises, chosen randomly from the 9 variations of the ring-transfer task, were performed in each session. The results showed a significant decrease in proficiency of performance in the postcall condition. In the postcall state, there was a 47% increase in cognitive errors across all exercises. However, the increase in errors was 56% in cognitively dominated exercises, showing that cognitive functioning is more significantly affected than psychomotor skills.

In this study, we employed the same apparatus and setup to test the impact of fatigue on neurosurgery residents. We recognized that the exercises do not in any way mimic actual neurosurgical procedures; at best, the measurement of cognitive and psychomotor skills is meant to parallel a measure of simulated surgical proficiency.

Methods

The study was conducted with the approval of Northwestern University’s institutional review board. The study participants were 7 neurosurgery residents in different years of training, ranging from PGY2 to PGY5, in the Northwestern University Feinberg School of Medicine neurological surgery residency program. Written informed consent was obtained from each resident prior to his or her participation. Baseline demographic data were collected from all 7 participants.

Simulation Exercises

We developed a series of virtual simulations that measured both psychomotor and cognitive skills in a controlled manner. A simulation was designed for the virtual ring-transfer task that is a part of a validated basic laparoscopic course using the ProMIS surgical simulator and FLS simulator (www.flsprogram.org). In the virtual ring-transfer task (Fig. 1), residents were tasked with grasping a series of “virtual” rings and placing each on randomly highlighted pegs on a board. The simulation was implemented using the Sensable haptic joystick, which allows for generation of 3 degrees of force feedback in response to events in the virtual environment. An OpenHL application programming interface was used to design the simulation. The simulation allows for measurement of the tool tip in the virtual environment. The basic task involves 10 rings. After the participant places a ring on a highlighted peg, another peg is randomly highlighted for the participant to put the ring on. This is repeated until all 10 rings are correctly placed. This basic ring-transfer task is a psychomotor task employed in many simulators to hone tool manipulation skills. A cognitive error is recorded every time the participant attempts to place a ring on the wrong peg (signifying error in judgment). It should be noted that the simulation does not allow placement of a ring on the wrong peg, and the participant is required to continue selecting pegs to put the ring on until the correct peg is chosen.

This basic validated laparoscopic exercise was modified to include cognitive variations. These cognitive variations were developed by employing neuropsychological methods that focus on developing tasks and exercises that measure cognitive abilities such as attention, visuospatial tracking, and intermodal transfer. Nine variations were designed, and these were described in detail by Kahol et al.,7 who validated the exercises as a tool to measure psychomotor and cognitive proficiency. Ring-transfer exercises are representative of laparoscopic tasks and are adequate for comparison of the performance of general surgeons and neurosurgeons.

Data Capture Tools and Proficiency Measurements

To assess fatigue levels of the participants, a questionnaire designed by Behrens and Monga1 was employed. Another questionnaire was developed for self-reporting of sleep obtained during call and the amount of caffeine consumption.

For measuring surgical proficiency, we employed a combination of time, cognitive errors, and tool movement. Tool-movement smoothness was measured on a scale of 0–1, with 0 indicating the least smoothness of movement and 1 the highest smoothness of movement. Tool movement, measured as movement of the tool tip in a virtual environment is a validated measure for surgical proficiency.4 In our simulations, tool-movement smoothness is determined by the simulator for an entire exercise. Senior surgeons tend to show a high degree of tool-movement
smoothness while novices tend to have lower values.\(^{10}\) In a simulation, the time required to complete a task is also recorded. The time elapsed can range from 0 to 300 seconds, but all raw data were normalized to a range of 0–1 and are thus reported as a percentage. Cognitive errors, defined as the number of times the ring was placed on the wrong peg, were also recorded for every type of exercise; once again, all raw data were normalized to a range of 0–1 and are thus reported as a percentage.

These 3 measures of proficiency (tool-movement smoothness, time elapsed, and cognitive errors) provided a broad framework for evaluation, and when coupled with fatigue and sleep-deprivation measures through the questionnaire, allow for holistic evaluation.

**Experimental Protocol**

We based our power analysis on previous studies of general surgeons. It was calculated that for a 2-tailed alpha of 0.05 with power of 80%, we needed a sample size of 5. For this study, we evaluated the pre- and postcall performance of 7 neurosurgeons.

For each participant, a session was defined as testing in the precall and corresponding postcall (sleep-deprived) state wherein “call” consisted of a 24-hour period of in-house call responsibilities. Participants completed a minimum of 3 and a maximum of 4 sessions. Only 1 session was allowed per week for each resident. Each testing session was preceded by the completion of the Behrenz questionnaire, which quantifies residents’ fatigue levels using self-reported assessments.

For each experiment, 3 exercises were randomly chosen from the 9 exercises available. Each exercise was repeated 2 times to counteract the warm-up effect in which the participant may perform suboptimally during the first exposure to a game. As with the previous methodology,\(^7\) the exercises in the precall condition and the postcall condition were not matched to account for a learning effect.

The designed software captured and stored the 3 proficiency measures. Fatigue ratings were captured through the questionnaire. Iterations of exercises performed pre-call were compared with iterations of the exercises performed postcall. An ANOVA was employed to study the difference between precall and postcall performance on each of the 3 proficiency measures of surgical skills. These measures enabled comparison of the effect of sleep deprivation on the surgical skills of the residents. Differences with probability values less than 0.05 were accepted as statistically significant.

**Results**

A total of 7 residents from PGY2 to PGY5 participated in the study; all completed 3 or 4 sessions. A session was defined as performance of the exercises in both the pre- and postcall states, with a call consisting of a 24-hour period of in-house call responsibilities. In total, data from 26 sessions were obtained and analyzed. The specific breakdown of the sessions is as follows: 2 PGY2 residents completed 4 sessions each; 1 PGY3 resident completed 3 sessions; 2 PGY4 residents completed 4 sessions each; 1 PGY5 resident completed 3 sessions; and 1 PGY5 resident completed 4 sessions (Table 1).

Analysis failed to show any statistically significant difference in residents’ surgical skills between the pre- and postcall states. The mean values for movement smoothness, time elapsed, and cognitive errors in the precall condition were 0.56, 0.45, and 0.34, respectively; in the postcall condition, the respective values were 0.5, 0.39, and 0.41, respectively (Fig. 2). The difference between pre- and postcall performance was not found to be statistically significant (p < 0.3, p < 0.4, and p < 0.2 for movement smoothness, time elapsed, and cognitive errors, respectively).

Taking movement smoothness and cognitive errors together as a measure of surgical proficiency, the residents in the postcall state exhibited a decrease in performance of 13.1%. This decrement was not found to be statistically significant. In contrast, in a similar study performed in general surgery residents, a statistically significant decrement of 27.3% was found in the postcall state. We did not find a significant correlation between cognitive errors and caffeine consumption (r = 0.45); this points to a limited effect of stimulants on performance.

**Discussion**

The effect of sleep deprivation on physicians has been actively researched over the past 2 decades; a number of studies have investigated residents’ ability to perform under conditions of sleep deprivation. Multiple studies have examined residents across different fields, quantifying the effects of sleep deprivation on both cognitive and psychomotor abilities.\(^2,3,5,6,9,11\) However, there has been a dearth of studies investigating the impact of fatigue on neurosurgery residents.

In our study, when comparing neurosurgery residents’ skills in the pre- and postcall states, no significant difference was found in the parameters of movement smoothness, elapsed time, and cognitive errors made. Specifically, neurosurgery residents, in the postcall state, demonstrated a 10.7% decrease in movement smoothness.

![Example of the virtual ring-transfer task. The highlighted peg is the one that the ring must be placed upon.](image-url)
Effect of call on neurosurgery residents’ skills

Table 1: Sessions completed by year of training

<table>
<thead>
<tr>
<th>Participant No.</th>
<th>Year of Residency</th>
<th>Sessions Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PGY2</td>
<td>x x x x</td>
</tr>
<tr>
<td>2</td>
<td>PGY2</td>
<td>x x x x</td>
</tr>
<tr>
<td>3</td>
<td>PGY3</td>
<td>x x x</td>
</tr>
<tr>
<td>4</td>
<td>PGY4</td>
<td>x x x x</td>
</tr>
<tr>
<td>5</td>
<td>PGY4</td>
<td>x x x</td>
</tr>
<tr>
<td>6</td>
<td>PGY5</td>
<td>x x x x</td>
</tr>
<tr>
<td>7</td>
<td>PGY5</td>
<td>x x</td>
</tr>
</tbody>
</table>

Fig. 2. Comparison of pre- and postcall performance in neurosurgeons and general surgeons.

and a 20.6% increase in cognitive errors; this was not statistically significant. In a similar study, general surgery residents showed a statistically significant decrease of 42.4% in movement smoothness and a 46.5% increase in cognitive errors in the postcall state. The results support our hypothesis that sleep deprivation does not affect physicians equally or uniformly. It is recognized by the researchers that the exercises do not in any way mimic actual neurosurgical procedures; at best, what we offer is an extrapolation in regard to simulated surgical proficiency. It is entirely possible that these results have no correlation with actual surgical performance in sleep-deprived states. Future work involves developing exercises that are more neurosurgery specific in terms of technical execution. In addition, further work may address the effect of sleep deprivation on clinical decision making in neurosurgery.

Graduate medical education underwent a major transformation in 2003 with the institution of work-hour restrictions; a major criticism of the policy is that it applies a “one size fits all” solution to a diverse group. Medicine is a complex and diverse field, and the subspecialties can and will make unique demands on their practitioners. The Institute of Medicine and ACGME should recognize the great variation in the medical specialties and their practitioners. The impact of fatigue on different specialties should be further investigated prior to implementation of a national physician work-hour policy.

An important element to consider is also not just the number of hours in the call period, but the type of duties a resident may perform. It is important to consider that certain tasks may be harder to perform when fatigued while routine tasks can be performed with acceptable accuracy. This study did not investigate this question in detail, but future work will include experiments to investigate this research question.

Conclusions

The results of our pilot study reveal that postcall fatigue is associated with a marginal decrease in surgical proficiency in neurosurgery residents. In contrast, in a comparative study, general surgery residents showed a statistically significant decrement of 27.3% in the postcall condition. We believe that the impact of fatigue on different specialties should be further investigated prior to implementation of a national physician work-hour policy.

The recent recommendations by the ACGME were based on a significant amount of effort, planning, and coordination. However, a systemic problem in the process of coming up with duty-hour recommendations is that while agencies invest a significant amount of effort, planning, and coordination, the scarcity of data on the impact of fatigue hampers the overall outcome. In the current era of technological advances, including electronic medical records, radiofrequency identification tags (employed to measure systemic errors and workflow), and surgical simulation (shown to have a positive impact on medical education by measuring the impact of fatigue), we believe that the committee should recommend using technologies to gather data on various aspects of this policy in the future. Without a coordinated plan and a program to employ these technologies, future ACGME policy and recommendations will again be based on sparse and insufficient data. This is clearly unacceptable from a safety and cost-effectiveness perspective. Agencies such as the National Institutes of Health, the Agency for Healthcare Research and Quality, ACGME, and the medical specialty boards should facilitate data collection and analysis on a large scale to study graduate medical education policy. We have the necessary tools and technologies to study and improve graduate medical education until the next iteration of review is scheduled. Practitioners of medicine should become proactive rather than reactive in the investigation and implementation of measures that lead to decreased medical errors and increased patient safety.

Disclosure

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author contributions to the study and manuscript preparation include the following. Conception and design: all authors. Acqui-
sition of data: Lee, Simonian, Quinn. Analysis and interpretation of data: Kahol. Drafting the article: Ganju, Kahol, Simonian. Critically revising the article: Ganju, Kahol, Lee, Simonian, Batjer. Statistical analysis: Kahol. Administrative/technical/material support: Kahol, Quinn, Batjer. Study supervision: Ganju, Simonian.

Acknowledgments

The authors would like to thank Linda Savitskas and Michael Johnson for their assistance in technological support and manuscript preparation.

References


Manuscript submitted August 20, 2010. Accepted September 25, 2011.

Portions of this work were presented as an oral abstract at the American Association of Neurological Surgeons meeting in Philadelphia, May 4, 2010.

Current affiliation for Dr. Kahol: Division of Affordable Health Technologies, Public Health Foundation of India, Delhi, India.

Please include this information when citing this paper; published online December 2, 2011; DOI: 10.3171/2011.9.JNS101406.

Address correspondence to: Aruna Ganju, M.D., Department of Neurological Surgery, Northwestern University Feinberg School of Medicine, 676 North St. Clair, Suite 2210, Chicago, Illinois 60611. email: aganju@nmff.org.