Wrist Motion Variation between Novices and Experienced Surgeons Performing Simulated Airway Surgery

Edward Callahan[1], Randall Bly, MD[1], Kaalan Johnson, MD[1],
Nava Aghdasi[2], Blake Hannaford, PhD[2], Kris Moe, MD[1], and
Maya G. Sardesai, MD, MEd[1]

Abstract

Objective. To determine whether wrist motion measured by a smartphone application can be used as a performance metric for a simulated airway procedure requiring both wrist and finger dexterity. We hypothesized that this accelerometer application could detect differences between novices and experienced surgeons performing simulated cricothyrotomy.

Setting. Academic medical center.

Study Design. Prospective pilot cohort study.

Methods. Voluntary surgeons and nonsurgeons were recruited. After viewing a training video, smartphones with accelerometer applications were attached to both wrists while subjects performed a cricothyrotomy on a validated task trainer. Procedure time and motion parameters, including average resultant acceleration (ARA), total resultant acceleration (TRA), and suprathreshold acceleration events (STAEs), were collected for dominant and nondominant hands. Subjects were stratified by prior experience. Blinded experts scored each performance using Objective Structured Assessment of Technical Skills (OSATS), and *t* tests were used to compare performance.

Results. Thirty subjects were enrolled. Median age was 26 years, and 20 subjects were male. In the dominant hand, significant differences were seen between novice and experienced surgeons in TRA ($P = .005$) and procedure time ($P = .006$), while no significant differences were seen in STAEs ($P = .42$) and ARA ($P = .33$). In the nondominant hand, all variables were significantly different between the 2 groups: STAEs ($P = .012$), ARA ($P = .007$), TRA ($P = .004$), and procedure time ($P = .006$).

Conclusions. Wrist motion measured by a low-cost smartphone application can distinguish between novice and experienced surgeons performing simulated airway surgery. This tool provides cost-effective and objective performance feedback.

Resident duty-hour restrictions have presented a challenge for educators and administrators in teaching hospitals. The impact of reduced procedural experience has implications on technical proficiency of trainees with fewer opportunities to train in the operating room and ethical concerns around residents practicing procedural skills on live patients. As a result, there has been an increasing interest in the use of simulation for technical skills education in surgery.[1-3]

The current standard for assessing trainee performance in simulated and in vivo settings is the Objective Structured Assessment of Technical Skill (OSATS),[4] which can be costly and time-consuming since it requires completion and interpretation by experts. Because of this, more cost-

[1]Department of Otolaryngology–Head and Neck Surgery, University of Washington School of Medicine, Seattle, Washington, USA
[2]Department of Engineering, University of Washington, Seattle, Washington, USA

Portions of this data were presented at the Western Medical Research Conference; January 28-30, 2016; Carmel, California.

Corresponding Author:
Maya G. Sardesai, MD, MEd, University of Washington School of Medicine Department of Otolaryngology–Head and Neck Surgery, Box 359803, 908 Jefferson St, Seattle, WA 98104, USA.
Email: sardesai@uw.edu

effective ways in which to provide feedback on performance during training are being explored. An example includes motion tracking using a hands-free device, which has been shown to demonstrate differences in parameters between novices and experts performing surgical procedures, including craniotomies and laparoscopic surgery.[5-11] This has not been well studied in otolaryngological procedures that may require greater digital dexterity than these other procedures. However, some vectors of wrist movement may be proxies for distal motion and may thus serve as valid metrics for performance.[10] Such metrics, if validated, could enable quantitative, objective assessment that may be less resource intensive than the current standard, which requires expert supervision.[4,12-14]

We hypothesize that wrist motion as measured by an accelerometer sensor on a smartphone will detect differences in performance between novices and experienced surgeons performing a cricothyroidotomy on a task trainer. Specifically, we predict that the total resultant acceleration (TRA) and average resultant acceleration (ARA) measured in both the dominant and nondominant wrists will correlate with procedural experience such that novices will demonstrate higher average and total values, as well as more frequent suprathreshold acceleration events (STAEs), denoting poorer fine motor control on both dominant and nondominant wrists. We also predict that experienced surgeons will perform the procedures more efficiently in less overall time. As such, wrist motion tracking could serve as an inexpensive method to provide objective real-time feedback to supplement surgical training.

Materials and Methods

Institutional review board approval was obtained through the University of Washington Human Subjects Division. Voluntary subjects, including undergraduates, medical students, residents, and otolaryngology faculty, were recruited using flyers and a designated script. After viewing a publicly available *New England Journal of Medicine* training video[15] demonstrating an example of the procedure, smartphones with accelerometer applications were fastened to both wrists while subjects performed a cricothyrotomy on a previously validated task trainer[3] designed in the engineering laboratory at the University of Washington. The model included a simulated airway with thyroid cartilage, cricoid cartilage, and tracheal rings covered by a synthetic skin analog composed of bicycle innertube (**Figure 1**).

The accelerometer data were collected using a commercially available application, Accelerometer Data Pro (Wavefront Labs, Austin, Texas). It was used to record acceleration in 3 axes at a frequency of 50 Hz. A filter mode was selected that adjusted for the effect of acceleration caused by gravity. The smartphones were attached firmly and painlessly to each subject's wrists using plastic wrap. The accelerometer applications used were in 2 Apple iPhones (Apple Corporation Cupertino, California). Based on resource availability, one was an iPhone 5 while the other was an iPhone 6, which were randomly placed on either hand. The 2 phones are identical in accelerometer sensor hardware (although the iPhone 6 does
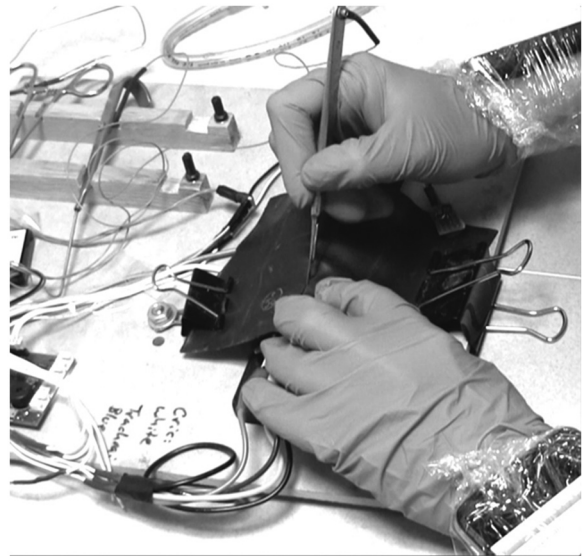

Figure 1. Photograph of assembly of wrist-mounted smartphones and cricothyrotomy simulator.

have an additional accelerometer that is used for low-power, less precise functions such as adjusting the screen orientation).

From the raw accelerometer data, multiple parameters were calculated, including the sum of resultant acceleration values (TRA) from the dominant and nondominant hands, the average resultant acceleration (ARA), and the number of STAEs, defined as greater than 1.3 m/s^2, and time to completion.[7] STAEs, as hypothesized by Jensen Ang et al,[7] represent a parameter that detects rapid and possibly uncontrolled movements. While they used an arbitrarily determined threshold of 5.8 m/s^2 to rule out any fine tremors, we reduced the threshold to 1.3 m/s^2 to increase sensitivity.

Subjects were stratified according to prior experience, where novices were defined as those with 0 to 15 similar prior procedures and the experts were those who had more than 15 similar prior procedures based on prior thresholds for plateaus of learning curves in other otolaryngology procedures.[16] "Similar procedures" were expected to include cricothyrotomies and tracheotomies in urgent and nonurgent settings. Primary outcome measures were TRA, ARA, STAEs, and total procedure time. Groups were compared using paired t tests with significance defined a priori at $P <$ .05. In addition, 2 blinded experts reviewed all video-recorded performances that did not include any specific identifying data and scored them using the OSATS. This included standardized scoring of 5 parameters: respect for tissue, time and motion, instrument handling, flow of operation, and knowledge of specific procedure, each on an anchored 5-point scale. Mean scores were calculated for each performance, and the mean score for each cohort was determined.

Results

Thirty subjects were recruited for the study. Eleven of the 20 novices had no prior experience, and 10 had experience with more than 15 prior similar procedures. Demographic

Table 1. Demographic Characteristics of Study Groups.

Characteristic	Novices (n = 20)	Experts (n = 10)
Median age, y	25	36
Sex, male, No.	15	5
Right hand dominant, No.	18	9
Mean OSATS score (P = .062)	15.0	19.1

Abbreviation: OSATS, Objective Structured Assessment of Technical Skills.

characteristics and mean OSATS score for each cohort are given in **Table 1**. The average time to complete the surgical task was 94.9 seconds, with a range from 30 seconds to 5 minutes.

Overall Time

Novices (fewer than 15 prior procedures) completed each task significantly slower than experts did at 106.4 seconds per trial, compared to 72 seconds for experts (P = .006). The overall time reached a plateau at around 16 prior similar procedures (**Figure 2**).

TRA

TRA was significantly different between novices and experts for both the dominant and nondominant wrist. The dominant wrist for the novices averaged 1081 m/s^2 while experts averaged 712 m/s^2 (P = .005). The nondominant wrist gave similar results; novices averaged 1041 m/s^2 and experts averaged 700 m/s^2 (P = .004) (**Figure 3**).

ARA

Only nondominant wrists showed a significant difference between expert and novice surgeons, with novices displaying an average of 0.984 m/s^2 and experts 0.973 m/s^2 (P = .007). The dominant wrists of novices demonstrated a mean ARA of 1.006 m/s^2 while experts had a mean ARA of 0.997 m/s^2 (P = .33) (**Figure 4**).

STAEs

STAEs were significantly different between novices and experts for the nondominant wrist. There was an average of 1.50

STAEs per procedure for experts and 4.75 STAEs for novices (P = .012). The dominant wrist, however, did not show a significant difference between experts and novices, with STAE average values of 10.8 and 11.6, respectively (P = .42) (**Figure 5**).

Discussion

This study identified small differences in multiple objective parameters recorded by a wrist accelerometer tracking novices and experts who performed a simulated cricothyroidotomy procedure. Specifically, small but statistically significant differences in TRA (in both wrists), ARA (nondominant wrist), and STAEs (nondominant wrist) were found. Although accelerometer data have not been recorded for cricothyroidotomy procedures in the past, our data are in accordance with prior work using accelerometers to track hand motion in laparoscopic surgery, microlaryngeal surgery, and neurosurgery.[5-11,17] The results are also in agreement with studies that have used more complicated motion tracking systems that are resource intensive and may not be able to passively record data.[12]

Jensen Ang et al[7] found that experts performed simulated neurosurgical procedures with lower TRA and less variability between dominant and nondominant hands. In this study, the statistically significant difference was only found in the nondominant wrist for ARA and STAEs. This supports the theory that function of the nondominant hand may be an important aspect of what constitutes an "expert performance."

The literature to support the claim that the nondominant hand is more important to stratify skill level is mixed. In 2 recent studies, the results did substantiate the importance of the nondominant hand in simulated orthopedic surgery[18] and vascular surgery.[19] However, other studies in laparoscopic surgery using microelectromechanical gyroscope tracking devices have found that differences are only statistically different for the dominant hand.[20] Despite this, many experienced surgeons believe gaining skill with the nondominant hand is critical for competency during surgery.[21-23]

In this study, we define experts as those who have performed greater than 15 similar procedures, and we assume that experts will carry out an "expert performance" each time. Similarly, a "novice performance" is defined as one performed by an individual with limited experience
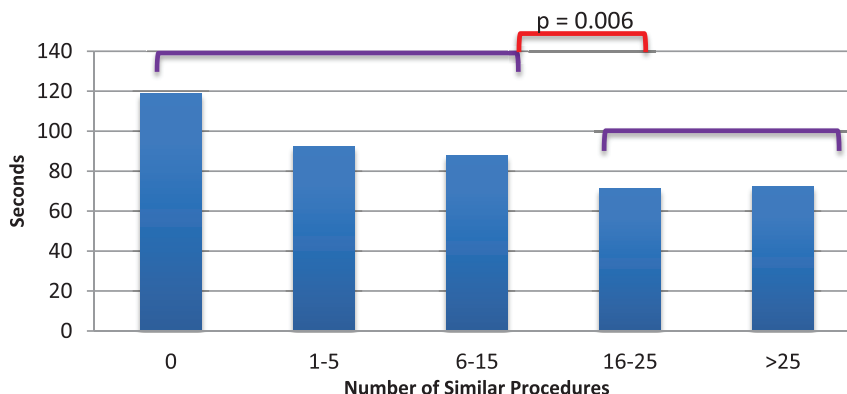

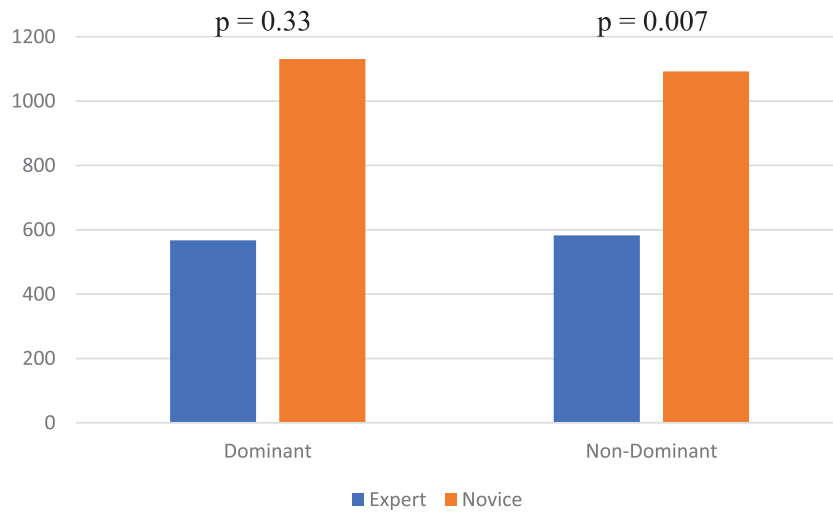**Figure 2.** Average overall procedure time for each experience level.

Figure 3. Total resultant acceleration comparing novices and experts in both dominant and nondominant wrists.
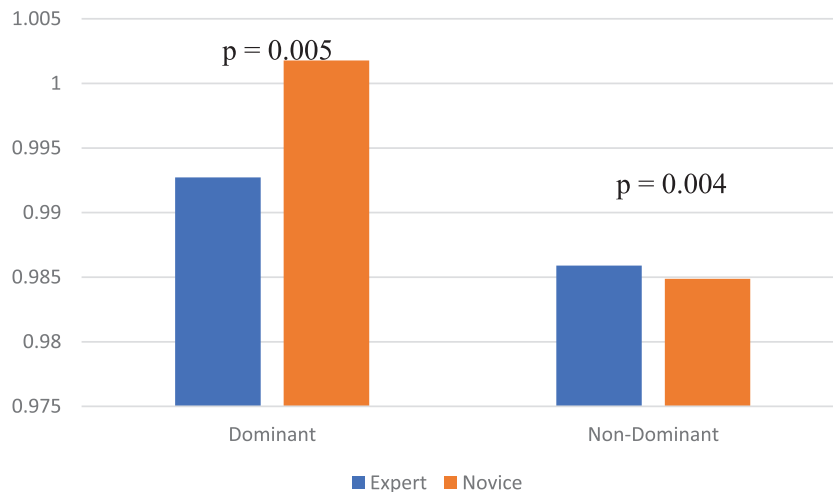

Figure 4. Average resultant acceleration comparing novices and experts in both dominant and nondominant wrists.

performing similar tasks. The skill level of the performances was evaluated using the OSATS, which is the current gold standard metric for residency trainees. Of note, although wrist motion varied significantly between novices and experts in the parameters identified above, the OSATS did not vary significantly between novice and experts in this study. This may in fact reflect a limitation to the use of OSATS in this setting with the possibility that wrist motion may detect more subtle differences in performance. The mean OSATS scores did trend toward a significant difference between novices and experts, however, and the absolute difference was small, which could help explain why the differences in wrist parameters were also small.

A limitation of this study is that the method was not validated by showing equivalence in tasks in which both cohorts can be assumed to be proficient, such as opening bottles and tying shoes. This is likewise a limitation of prior similar studies[7] and can be addressed in subsequent work. Another limitation to the interpretation of the data is that

the performances were done on a task trainer without the stress of urgent airway surgery in a real patient. A such, the transfer of performance on the task trainer to the in vivo setting has yet to be determined.

The demographic data were limited by self-reporting, which may have carried some recall bias. Furthermore, this procedure represents a high-acuity, low-frequency event, and as such, even experienced surgeons may not have had recent applicable experiences. Finally, the small sample size and homogeneous population limit the generalizability of the data.

An additional question is whether acceleration is in fact an appropriate metric of performance. Arguably, an aspect of surgical expertise is knowing when it is most appropriate to use slow and careful movements and when increased speed is safe and efficient. This may influence changes seen in the acceleration. An evaluation that incorporates degree of tissue trauma and risk of complications could be of benefit. Regardless, acceleration, even at a proximal level, does clearly accurately stratify skill level in this study, as is substantiated in the literature.
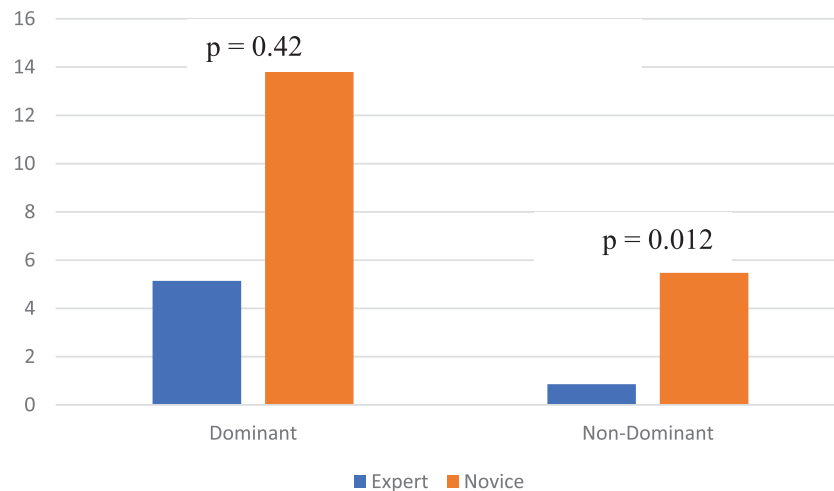
Figure 5. Suprathreshold acceleration events comparing novices and experts in both dominant (*P* = .42) and nondominant wrists (*P* = .012).

Feedback for surgical trainees is more important now than ever due to reduced work hours, possibly less experience in the operating room during training, and more demands from credentialing societies to demonstrate competency. Useful feedback can be challenging to provide due to the high cost of expert surgeon time.[3] Furthermore, when this feedback is provided, it is rarely objective. The current standard is a standardized form (OSATS) completed by a supervising surgeon.[4] This provides feedback in multiple domains related to technical skill, but it is time intensive to the expert surgeon, and although structured, it is not always objective.

The results in this study represent a major opportunity to supplement standard subjective feedback that is provided to trainees, and it does so passively with little expense. The implications are that a surgical trainee could wear a wristwatch fitted with an accelerometer, and his or her motions could be tracked over time. There is likely to be a threshold of accelerometer-derived parameters that correlate with that individual's attainment of surgical competency. There are many opportunities for future study to answer the question of how many procedures are required to attain competence in a given procedure. This method would allow that question to be answered quantitatively, as it is compared with current evaluation methods.

Conclusions

Wrist motion measured by a low-cost smartphone application varies slightly but significantly between novice and experienced surgeons performing simulated airway surgery, particularly in the nondominant hand. With further study and refinement, this tool has the potential to provide a cost-effective real-time objective supplement to surgical education and feedback.

Author Contributions

Edward Callahan, data collection, analysis, manuscript preparation, revision, approval, accountability; **Randall Bly**, data analysis, manuscript revision, approval, accountability; **Kaalan Johnson**, data analysis, manuscript revision, approval, accountability; **Nava**

Aghdasi, data collection, analysis, manuscript revision, approval, accountability; **Blake Hannaford**, data analysis, manuscript revision, approval, accountability; **Kris Moe**, design, data analysis, manuscript revision, approval, accountability; **Maya G. Sardesai**, design, analysis, manuscript preparation, revision, approval, accountability.

References

1. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg.* 2014;259:1041-1053.
2. Gawande A. The learning curve: like everyone else, surgeons need practice. That's where you come in. *New Yorker.* http://www.newyorker.com/magazine/2002/01/28/the-learning-curve. Accessed April 12, 2017.
3. Aghdasi N, Bly R, White LW, et al. Crowd-sourced assessment of surgical skills in cricothyrotomy procedure. *J Surg Res.* 2015;196:302-306.
4. Sultana CJ. The objective structured assessment of technical skills and the ACGME competencies. *Obstet Gynecol Clin North Am.* 2006;33:259-265.
5. Ghobrial GM, Balsara K, Maulucci CM, et al. Simulation training curricula for neurosurgical residents: cervical foraminotomy and durotomy repair modules. *World Neurosurg.* 2015;84:751-755.
6. Jordan A, Antomarchi J, Bongain A, et al. Development and validation of an objective structured assessment of technical skill tool for the practice of breech presentation delivery. *Arch Gynecol Obstet.* 2016;294:327-332.
7. Jensen Ang WJ, Hopkins ME, Partridge R, et al. Validating the use of smartphone-based accelerometers for performance

assessment in a simulated neurosurgical task. *Neurosurgery*. 2014;10(suppl 1):57-65.

8. Hofstad EF, Våpenstad C, Chmarra MK, et al. A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance. *Surg Endosc*. 2013;27:854-863.

9. Katsavelis D, Siu KC, Brown-Clerk B, et al. Validated robotic laparoscopic surgical training in a virtual-reality environment. *Surg Endosc*. 2009;23:66-73.

10. Oropesa I, Chmarra MK, Sánchez-González P, et al. Relevance of motion-related assessment metrics in laparoscopic surgery. *Surg Innov*. 2013;20:299-312.

11. Nagendran M, Toon CD, Davidson BR, et al. Laparoscopic surgical box model training for surgical trainees with no prior laparoscopic experience. *Cochrane Database Syst Rev*. 2014;(1):CD010479.

12. Conroy E, Surender K, Geng Z, et al. Video-based method of quantifying performance and instrument motion during simulated phonosurgery. *Laryngoscope*. 2014;124:2332-2337.

13. Zevin B, Aggarwal R, Grantcharov TP. Surgical simulation in 2013: why is it still not the standard in surgical training? *J Am Coll Surg*. 2014;218:294-301.

14. Tsue TT, Dugan JW, Burkey B. Assessment of surgical competency. *Otolaryngol Clin North Am*. 2007;40:1237-1259.

15. Hsiao J, Pacheco-Fowler V. Videos in clinical medicine: cricothyroidotomy. *N Engl J Med*. 2008;358:e25.

16. Laeeq K, Pandian V, Skinner M, et al. Learning curve for competency in flexible laryngoscopy. *Laryngoscope*. 2010;120:1950-1953.

17. Chen T, Surender K, Vamos AC, et al. Quantitative evaluation of phonomicrosurgical manipulations using a magnetic motion tracking system. *Laryngoscope*. 2014;124:2107-2113.

18. Pedowitz R, Nicandri G, Tuchschmid S. Asymmetry in dominant/non-dominant hand performance differentiates novices from experts on an arthroscopy virtual reality serious game. *Stud Health Technol Inform*. 2016;220:289-294.

19. Bech B, Lönn L, Schroeder TV, et al. Fine-motor skills testing and prediction of endovascular performance. *Acta Radiol*. 2013;54:1165-1174.

20. Overby DW, Watson RA. Hand motion patterns of Fundamentals of Laparoscopic Surgery certified and noncertified surgeons. *Am J Surg*. 2014;207:226-230.

21. Papaspyros SC, Kar A, O'Regan D. Surgical ergonomics: analysis of technical skills, simulation models and assessment methods. *Int J Surg*. 2015;18:83-87.

22. Rosser JC, Rosser LE, Savalgi RS. Skill acquisition and assessment for laparoscopic surgery. *Arch Surg*. 1997;132:200-204.

23. Gupta R, Guillonneau B, Cathelineau X, et al. In vitro training program to improve ambidextrous skill and reduce physical fatigue during laparoscopic surgery: preliminary experience. *J Endourol*. 2003;17:323-325.