Objective Measurement of Clinical Competency in Surgical Education Using Electrodermal Activity

Jacob A. Quick, MD, Alex D. Bukoski, DVM, PhD,† Jennifer Doty, BSN, RN,* Bethany J. Bennett, MHA,* Megan Crane, LPN,* and Stephen L. Barnes, MD, FACS**

*Department of Surgery, School of Medicine, University of Missouri, Columbia, Missouri; and †Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri, Columbia, Missouri

OBJECTIVE: Within the realm of surgical education, there is a need for objective means to determine surgical competence and resident readiness to operate independently. We propose a novel, objective method of assessing resident confidence and clinical competence based on measurement of electrodermal activity (EDA) during live surgical procedures. We hypothesized that with progressive training, EDA responses to the stress of performing surgery would exhibit decline, elucidating an objective correlate of clinical competence.

DESIGN: EDA was measured using galvanic skin response sensors worn by residents performing laparoscopic cholecystectomy on sequential live human patients over an 8-month period. Baseline, phasic (peak) and tonic EDA responses were measured as a fractional change from baseline.

SETTING: University of Missouri, Columbia, Missouri, an academic tertiary care facility.

PARTICIPANTS: Fourteen categorical general surgery residents and 5 faculty surgeons were voluntarily enrolled and participated through completion.

RESULTS: Tonic fractional change (FC_{TONIC}) was highest in PGY3 residents compared with postgraduate year (PGY) 1 and 2 residents (7.199 vs. 2.100, $p = 0.004$, 95% CI: 8.58-1.61 and PGY4 and 5 residents (7.199 vs. 2.079, $p = 0.002$, 95% CI: 8.38-0.29). Phasic fractional change in EDA (FC_{PHASIC}) exhibited a progressive decline across resident training levels, with PGY1 and 2 residents having the highest response, and faculty displaying the lowest FC_{PHASIC} responses. Statistical differences were seen between FC_{PHASIC} faculty and PGY4 and 5 (3.596 vs. 6.180, $p = 0.004$, 95% CI: 0.80-4.36), PGY4 and 5, and PGY3 (6.180 vs. 15.998, $p = 0.003$, 95% CI: 3.33-16.3), as well as among all residents and faculty (13.057 vs. 3.596, $p = 0.004$, 95% CI: 15.8-3.1).

CONCLUSION: Phasic EDA changes decrease with increasing clinical competence. For those participants with the lowest and highest levels of competence, tonic EDA changes are minimal. Tonic EDA changes follow an inverse-U shape with differing levels of clinical competence. (J Surg Ed ▮:▮▮▮-▮▮▮. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: surgical education, electrodermal activity, galvanic skin response, competency, objective

COMPETENCY: Practice-based Learning and Improvement

INTRODUCTION

Surgical knowledge, measured via the American Board of Surgery Qualifying Examination, has steadily increased since 1980, whereas results from the Certifying Examination have declined slightly since initiation of work-hour restrictions.[1,2] Concern over clinical competency, confidence, and decision-making skills has thus been a point of interest for surgical residents, education programs, societies, and policy makers alike. Specific focus on resident readiness to operate independently is emphasized by several reports in recent years highlighting the relative lack of confidence in graduating residents. Transition to practice fellowships has been created to help improve the confidence and competence of the approximately one-quarter of graduating residents who believe that they are not ready to practice independently.[3,4] Gradual increases in medicolegal restrictions and public concern have further limited the

Correspondence: Inquiries to Jacob A. Quick, MD, Department of Surgery, School of Medicine, University of Missouri, 1 Hospital Drive MC220, Columbia, MO 65212; (573) 884-2835; e-mail: quickja@health.missouri.edu

  1931-7204/$30.00 http://dx.doi.org/10.1016/j.jsurg.2017.01.007

ability of residents to complete cases without direct faculty supervision. These findings highlight concerns over resident readiness to operate, and bring into question the educational methods and evaluation tools currently used to assess trainees' ability to function independently.

Current evaluation methods, although more complex, have not fundamentally changed since the days of Halsted and Osler. Subjective evaluations continue to permeate training programs and rely mainly upon the "gut feelings" of evaluators. Several systems have been developed to address the paucity of objective performance data, including the Objective Structured Assessment of Technical Skills (OSATS), which uses a combination of task-specific checklists (Yes/No) and global rating scales.[5,6] Although the task-specific portion provides objective measurement in black-and-white terms, global assessment remains entirely subjective. Variations and modifications of the OSATS model exist, but none are completely objective.[7] Other evaluation methods, such as the Ottawa Surgical Competency Operating Room Evaluation (O-SCORE) and Global Assessment of Gastrointestinal Endoscopic Skills (GAGES), although validated, are also wholly subjective.[8,9] The most recent Next Accreditation System, also known as the Milestones Project, from the Accreditation Council for Graduate Medical Education (ACGME) focuses on complex competency evaluation, however, remains largely subjective.[10]

Purely objective determination of resident competence has long relied upon resident case logs, with the underlying premise being that, "more cases, equals more competence." However, case logs do not provide insight into the actual abilities of a trainee to complete an operation independently. With up to one-third of general surgeons in fellowship training being unable to independently complete a laparoscopic cholecystectomy, there is evidence that resident operative logs merely offer details of the breadth and type of surgical exposure during training.[11] Several methods to objectively quantify competency are being investigated. Heart rate variability, which decreases during times of stress and has been used in psychological research for many years, remains in its infancy with regard to competence evaluation.[12-14] More sophisticated ways to objectively evaluate skill acquisition, including functional magnetic resonance imaging, are also being investigated, but they lack the ability to determine skill and competence during real-life surgical procedures because of sterility and operating room limitations.[15]

Electrodermal activity (EDA), as a measure of stress and conversely, comfort, is not a new concept. First observed over 150 years ago, EDA has been heavily used in psychophysiology research. Its use involves noninvasive dermal sensors to determine changes in the electrical conductance of the skin in response to stress. Changes in conductivity have been well described as a method for quantifying the level of stress and comfort during mentally taxing events.[16-18] Using EDA principles, a baseline level of skin conductance, or galvanic skin response, is recorded preceding a specified event. Once the event begins, 2 forms of EDA phenomena are logged: tonic level and phasic responses. Tonic levels represent the overall galvanic skin response during a defined period, and thus serve to quantify the overall level of engagement during the entire event or recording period. For example, the time between operation start and end would represent the tonic level. Phasic responses consist of fast upstrokes in EDA and occur at times of peak stress during an event. For example, when placing clips and dividing critical structures during laparoscopic cholecystectomy.

By measuring stress throughout the course of an operation, analysis of these 2 EDA stress responses can be interpreted by combining several psychological constructs, including the Yerkes-Dodson model and the Dunning-Kruger effect.[19-21] We hypothesized that as competence increases beyond the beginning stages, the level of stress experienced by the learner decreases. When mastery is attained, comfort with the task is at its maximum and stress at a minimum. We propose a novel, objective method to determine surgical resident comfort, confidence, and clinical competence using EDA. We hypothesized that with progressive training, EDA responses would decrease, signifying an objective index of increasing clinical competence.

METHODS

Following Institutional Review Board's (IRB) approval, voluntary informed consent was obtained from categorical general surgery residents (postgraduate years [PGY] 1 to 5) and faculty before participation in the study. Study results were blinded to the residency program, and no formal, program-sponsored resident evaluations were completed based upon the results. A waiver of consent for patients undergoing laparoscopic cholecystectomy during EDA evaluation was obtained from the IRB. Laparoscopic cholecystectomy was chosen as the study environment given the high frequency of the procedure, well-defined steps necessary for completion, and standardized case difficulty grading system availability. Operative cases were graded on a score of 1 to 5, with clear definitions outlined before and throughout the study. Scoring was based upon a modified Cuschieri scale.[22] Faculty observers underwent training sessions to ensure accurate case difficulty scoring using video review of surgical cases. Faculty observer concordance ($>80\%$) was ensured through repeated observer training sessions. Experienced operative assistance, either from faculty members or senior residents, was guaranteed throughout the study to minimize confounders related to inexperienced assistants. Sequential live human laparoscopic cholecystectomy operative cases during an 8-month period were included in the study for analysis. Operations