Surgery Tutor for Computational Assessment of Technical Proficiency in Soft-Tissue Tumor Resection in a Simulated Setting

Caitlin T. Yeo, MD, Justine Ring, BScH,[1] Matthew S. Holden, MSc, Tamas Ungi, MD, PhD, Ayca Toprak, MEd, FRCSC, Gabor Fichtinger, PhD, and Boris Zevin, PhD, FRCSC

Department of Surgery, Queen's University, Kingston, Ontario, Canada

BACKGROUND: In competency-based medical education, progression between milestones requires reliable and valid methods of assessment. Surgery Tutor is an open-source motion tracking platform developed to objectively assess technical proficiency during open soft-tissue tumor resections in a simulated setting. The objective of our study was to provide evidence in support of construct validity of the scores obtained by Surgery Tutor. We hypothesized that Surgery Tutor would discriminate between novice, intermediate, and experienced operators.

METHODS: Thirty participants were assigned to novice, intermediate, or experienced groups, based on the number of prior soft-tissue resections performed. Each participant resected 2 palpable and 2 nonpalpable lesions from a soft-tissue phantom. Surgery Tutor was used to track hand and instrument motions, number of tumor breaches, and time to perform each resection. Mass of excised specimens and margin status were also recorded.

RESULTS: Surgery Tutor scores demonstrated "moderate" to "good" internal structure (test-retest reliability) for novice, intermediate, and experienced groups (interclass correlation coefficient = 0.596, 0.569, 0.737; $p < 0.001$). Evidence in support of construct validity (consequences) was demonstrated by comparing scores of novice, intermediate, and experienced participants for number of hand and instrument motions (690 ± 190, 597 ± 169, 469 ± 110; $p < 0.001$), number of tumor breaches (29 ± 34, 16 ± 11, 9 ± 6; $p < 0.001$), time per resection (677 ± 331 seconds, 561 ± 210 seconds, 449 ± 148 seconds; $p < 0.001$), mass of completely excised specimens (22 ± 7 g, 21 ± 11 g, 17 ± 6 g; $p = 0.035$), and rate of positive margin (68%, 50%, 28%; $p < 0.001$). There was "strong" and "moderate" relationships between motion scores and Objective Structured Assessment of Technical Skill scores, and time per resection and Objective Structured Assessment of Technical Skill scores respectively ($r = -0.60$, $p < 0.001$; $r = -0.54$, $p < 0.001$).

CONCLUSION: Surgery Tutor scores demonstrate evidence of construct validity with regards to good internal structure, consequences, and relationship to other variables in the assessment of technical proficiency during open soft-tissue tumor resections in a simulated setting. Utilization of Surgery Tutor can provide formative feedback and objective assessment of surgical proficiency in a simulated setting. (J Surg Ed 76:872−880. © 2018 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: objective assessment, computer-based assessment, medical education, surgical education, technical skills

COMPETENCIES: Practice-Based Learning and Improvement

INTRODUCTION

Postgraduate medical education is undergoing a major shift toward Competency-Based Medical Education (CBME).[1,2] This transition to CBME has resulted in tremendous scholarly activity in the fields of curriculum

Financial support: This project was supported by the Queen's University Research Initiation Grant with grant number 389139.

Correspondence: Inquiries to Caitlin T. Yeo, MD Queen's University, Department of Surgery 76 Stuart Street, Victory 3 Kingston, ON K7L 2V7, Canada; e-mail: cyeo@qmed.ca

[1] Co-first author.

  1931-7204/$30.00
https://doi.org/10.1016/j.jsurg.2018.11.005

development, implementation, assessment, and evaluation. Timely transition between milestones within a CBME curriculum requires reliable, valid, and objective assessments of clinical competence at each training level. Within the discipline of surgery, multiple assessment tools have been developed for both technical and nontechnical skills. These include procedure-specific checklists,[3,4] global rating scales,[5-7] and motion analysis platforms.[8-10] Each assessment tool has specific advantages and disadvantages which guide its utilization.[3,11-13] Each tool individually does not provide adequate global assessment, but rather when used in combination can provide useful feedback on overall trainee ability and progression—a method known as triangulation.[14] A number of these tools have been incorporated into specific training curricula as a means of formative and summative assessment.[15,16]

Motion analysis platforms allow for the objective assessment of technical proficiency.[17] Technical proficiency is often defined as an adequate level of competence at a skill through learning and deliberate practice.[18] Proficiency requirements may evolve as a trainee progresses through their training. These devices rely on motion tracking hardware and software to detect the motion of an operator's hands or instruments. Examples of these platforms include the Imperial College Surgical Assessment Device (Imperial College, London, UK), the Advanced Dundee Psychomotor Tester (University of Dundee, Dundee, UK), and the Perk Tutor (Queen's University, Ontario, Canada). Advantages of motion tracking include the objectivity of metrics obtained by the device, autonomous assessment without the presence of a trained preceptor, and the ability to use these devices in both a simulation laboratory and in the operating room. The disadvantages of motion tracking devices include the requirement for hardware and software resources, the difficulty of translating the metrics of distance, speed, and time into a formative assessment, and the limitations of technology.

Training within a simulation laboratory has been shown to result in improved technical and nontechnical skills in a simulated setting and in the operating room.[17,19,20] At present, there are no valid objective means to assess technical proficiency of trainees and practicing surgeons during open excisions of soft-tissue tumors such as breast lumpectomy and soft tissue sarcoma. Given the competing requirements for surgeons to preserve surrounding healthy tissue and to obtain negative oncologic margins, improved technical proficiency during excision of soft-tissue tumors may result in improved oncologic outcomes, as well as improved function and cosmesis.

The objectives of our study were (1) to develop an open source Surgery Tutor platform for objective assessment of technical proficiency during open excision of soft-tissue tumors, and (2) to provide evidence in support of construct validity of the scores obtained by Surgery Tutor in a simulated setting.

METHODS

Study Design

This was a prospective cross-sectional study conducted in 2017 at Queen's University, Ontario, Canada. The study was approved by the University Health Science and Affiliated Teaching Hospitals Research Ethics Board. Voluntary enrollment and signed consent was obtained from each participant.

Study Setting and Population

Study participants included senior medical students, general surgery residents, and practicing general surgeons at Queen's University. The study was conducted at the Queen's University Clinical Simulation Centre. Study participants were assigned to a novice, intermediate, or experienced group based on the number of soft-tissue tumor resections they performed as the primary operator within the last 5 years prior to time of study enrollment. The novice group was defined as participants that had performed ten or fewer soft-tissue resections, the intermediate group was defined as those that had performed between 11 and 39 soft-tissue resections, and the experienced group was defined as those that had performed greater or equal to 40 soft-tissue resections.[21]

Study Protocol and Data Collection

Demographic information regarding study participant's level of training, number, and comfort level with soft-tissue and breast resections was collected. Each study participant watched a short video on how to perform an excision of a palpable and a nonpalpable soft-tissue tumor from a synthetic phantom to standardize the workflow, which consisted of surgical planning, excision, and closure (Fig. 1). Each participant was asked to resect 2 palpable and 2 nonpalpable tumors from the phantom. Participants were block randomized within their assigned groups to start with either the 2 palpable or the 2 nonpalpable tumors.

Surgery Tutor was used to record and compute metrics related to hand motion, instrument motion, tumor location, number of times scalpel breached the tumor, and procedural time (total time, and time spent on each phase of workflow). Mass of excised specimen and margin status (positive or negative) were also recorded.

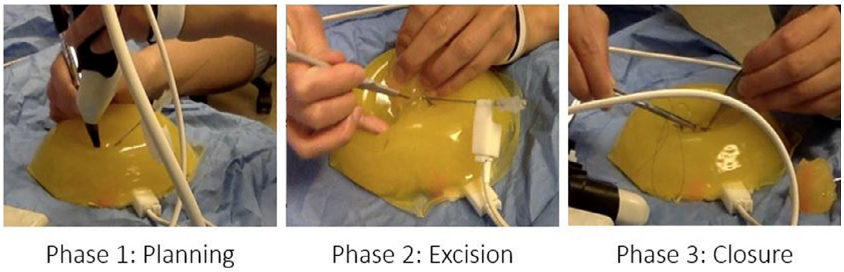

FIGURE 1. Screenshots of video demonstrating procedure workflow.

Participants were not able to view the metrics or the navigation display (Fig. 2).

All sessions were video recorded, ensuring that only the participant's hands and the setup were recorded without sound or any other identifying features to ensure blinded video assessment of operative performance. An independent trained reviewer with experience in soft-tissue resections and assessment of procedural skills reviewed the videos in a blinded fashion and assessed operative performance using the Objective Structured Assessment of Technical Skill (OSATS)[22] (Appendix, Table A). The maximum achievable score on the OSATS scale is 20.

Surgery Tutor Platform

Surgery Tutor is a novel platform that integrates 2 previously well tested systems, Perk Tutor[23-25] and Naviknife.[26] These systems are available through the open source 3D Slicer medical image analysis platform (www.slicer.org). The hardware consisted of an electromagnetic (EM) field generator and control box (Ascension, Milton, VT), and EM tracked sensors (Model 800, 7.9 mm, 6-DOF). The EM sensors track path length and translational and rotational motion of participant's hands and instruments (scalpel and ultrasound). The hand sensors were attached to the back of the operator's hands using soft Velcro straps. The tumor was marked with an EM tracked hooked guidewire and the tumor location was registered using ultrasound. Using the registered tumor location, and the platform is able to identify when the tumor is breached by the scalpel. Three-dimensional position data from the EM sensors were acquired using PLUS open source software (www.PlusToolkit.org).

A simulated breast phantom with palpable and nonpalpable tumors was used for our study. The phantoms were constructed from plastisol liquid plastic using a

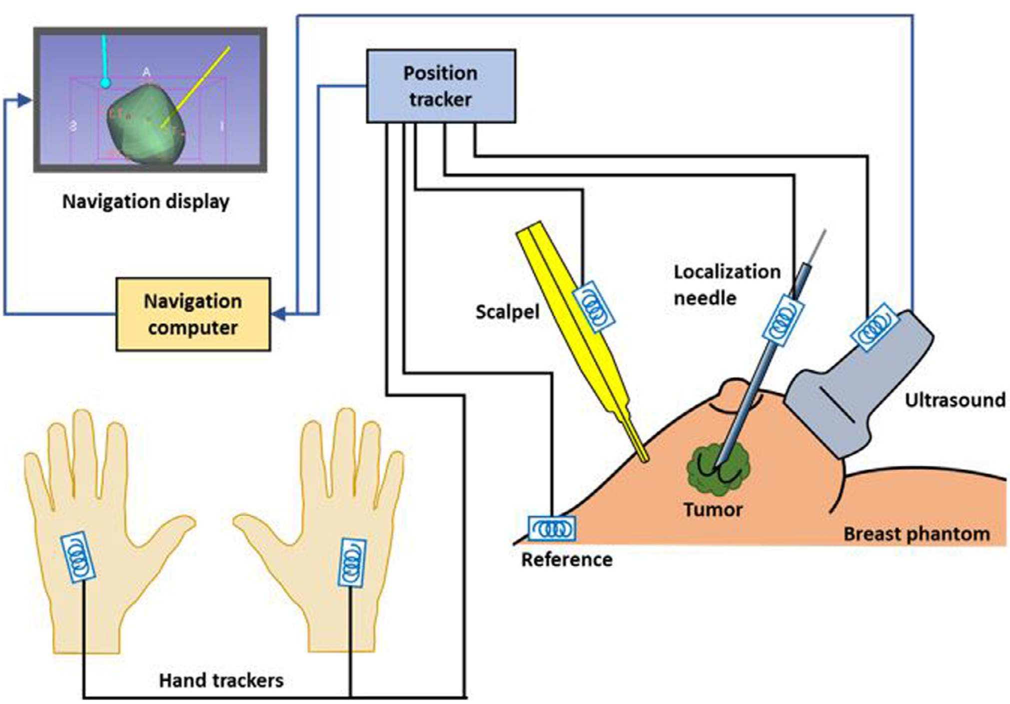

FIGURE 2. System setup demonstrating position of EM sensors and registered tumor location.

tested recipe that allows for realistic ultrasound imaging and tissue handling.[26]

Sample Size

Sample size was calculated based on previously published hand motion analysis data using the Perk Tutor platform.[25] We used a Cohen's $d = 0.5$ (medium) effect size, 2-tailed alpha = 0.05, power of 80%, and 1:1:1 assignment of individuals within groups (novice, intermediate, and experienced). A sample size of ten participants per study group was required to detect a statistically significant difference.

Construct Validity of Scores from Surgery Tutor

We utilized the unitary framework for construct validity proposed by Messick.[27] In this framework, evidence to support the construct validity argument is collected from 5 sources: content, response process, internal structure, relations to other variables, and consequences. For our study we focused on gathering evidence about internal structure, consequences of scores obtained from Surgery Tutor, and relations to other variables. Evidence about the internal structure was gathered from test-retest reliability of Surgery Tutor scores calculated using the interclass correlation coefficient (ICC) for the motion metrics of the first 2 consecutive tumor resections performed by the same participant. Evidence about consequences was gathered by comparing Surgery Tutor scores (metrics for hand and instrument motion, number of tumor breaches, procedure time, mass of completely excised specimens with negative margins only, and positive margin rate) between novice, intermediate, and experienced groups using the Kruskal-Wallis test for nonparametric data. Evidence about relations to other variables was gathered by comparing Surgery Tutor metrics (hand and instrument motion, time, and number of tumor breaches) against the OSATS scores using Spearman correlation coefficient for nonparametric data and Pearson correlation coefficient for parametric data.

p Values were adjusted for the comparison between 3 groups using the Bonferroni correction. Adjusted statistical significance was $p < 0.017$. Data is presented as median (range) for nonparametric analysis or mean ± standard deviation (SD) for parametric analysis. Statistical analysis was performed using IBM SPSS Statistics Version 24.

RESULTS

A total of 30 participants were recruited, with ten novice, ten intermediate, and ten experienced participants (Table 1). The median number of soft-tissue resections performed by participants in each group was 0 (0-6), 21 (11-40), and 88.5 (43-1000), respectively. Initial median level of comfort rated from 0 to 10 with soft-tissue resections (1.5 (0-8), 5 (3-7), 8.5 (7-10); $p < 0.001$) and breast lumpectomy (1.5 (0-5), 5 (2-8), 8(6-10); $p < 0.001$) was significantly different between groups (Table 1).

Evidence about the internal structure from test-retest reliability of Surgery Tutor scores demonstrated "moderate" to "good" reliability for novice group (ICC = 0.60, 0.01-0.86; $p < 0.001$), intermediate group (ICC = 0.57, 0.08-0.86; $p < 0.001$), and experienced group (ICC = 0.74, 0.40-0.90; $p < 0.001$) (Table 2).

Evidence about consequences was gathered by comparing Surgery Tutor scores between novice, intermediate, and experienced groups. There was a significant difference between novice, intermediate, and experienced groups for the number of hand and instrument motions (690 ± 190, 597 ± 169, 469 ± 110; $p < 0.001$), average number of tumor breaches per excision (29 ± 34, 16 ± 11, 9 ± 6; $p < 0.001$), total procedure time (677 ± 331 seconds, 561 ± 210 seconds, 449 ± 148 seconds; $p < 0.001$), and rate of positive

TABLE 1. Demographic Data; Median (Range)

	Novice (n = 10)	Intermediate (n = 10)	Experienced (n = 10)	P
Gender				
• M	5	7	5	
• F	5	3	5	
Level of training				
• Senior medical student	6	0	0	
• Junior resident (PGY 1-2)	4	4	0	
• Senior resident (PGY 3-5)	0	5	5	
• General surgeon	0	1	5	
Number of soft tissue resections	0 (0-6)	21 (11-40)	88.5 (43-1000)	<0.001
Number of breast lumpectomies	0 (0-4)	10 (5-15)	59 (0-500)	<0.001
Comfort with soft tissue resections (0-10)	1.5 (0-8)	5 (3-7)	8.5 (7-10)	<0.001
Comfort with breast lumpectomy (0-10)	1.5 (0-5)	5 (2-8)	8 (6-10)	<0.001

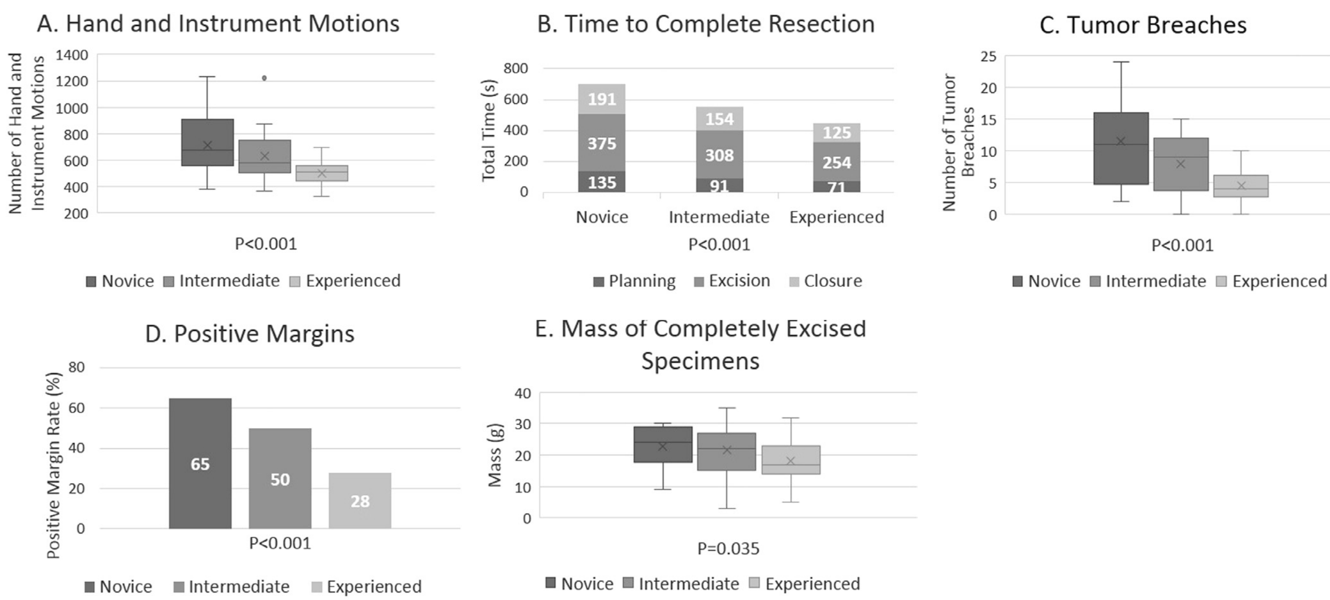

FIGURE 3. A. Number of hand and instrument motions, B. Time to complete resection, C. Number of tumor breaches, D. Positive margin rate, E. Mass of completely excised specimens.

margins (68%, 50%, 28%; $p < 0.001$)(Fig. 3A-D).There was a trend toward significant difference in the mass of completely excised specimens (22 g (11-30 g), 21g (3-53)g, 17 g (5-32)g; $p = 0.035$) (Fig. 3E).

Evidence about relations to other variables was gathered by comparing Surgery Tutor scores (hand and instrument motion, time, and number of tumor breaches) against OSATS scores. There was a significant difference in OSATS scores between novice, intermediate, and experienced groups (11 (7-13), 12 (10-15), 14.5 (11-18); $p < 0.001$) (Fig. 4). There was a "moderate" to "strong" correlation between hand/instrument motion and OSATS scores ($r = -0.60$, $(-0.81) - (-0.27)$; $p < 0.001$), and between procedural time and OSATS scores ($r = -0.54$, $(-0.76) - (-0.20)$; $p < 0.001$) (Fig. 5a-b). There was a weak correlation between the number of tumor breaches and OSATS scores ($r = -0.33$, $(-0.63) - (0.05)$; $p = 0.093$) (Fig. 5c).

TABLE 2. Test-Retest Reliability

	ICC	P
Novice	0.60 (0.01-0.86)	p<0.001
Intermediate	0.57 (0.08-0.86)	p<0.001
Experienced	0.74 (0.40-0.90)	p<0.001

DISCUSSION

Our results demonstrate the construct validity of the scores obtained by Surgery Tutor for assessment of technical proficiency in soft-tissue tumor resections in a simulated setting. We demonstrated a "moderate" to "good" test-retest reliability of scores for novice, intermediate, and experienced groups, providing evidence of internal structure. Hand and instrument motion, number of tumor breaches, time per procedure, mass of completely excised specimens, and positive margin rates were significantly different between participants in novice,

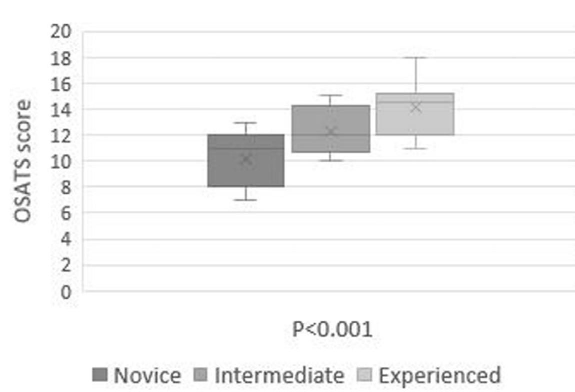

FIGURE 4. OSATS scores by group.

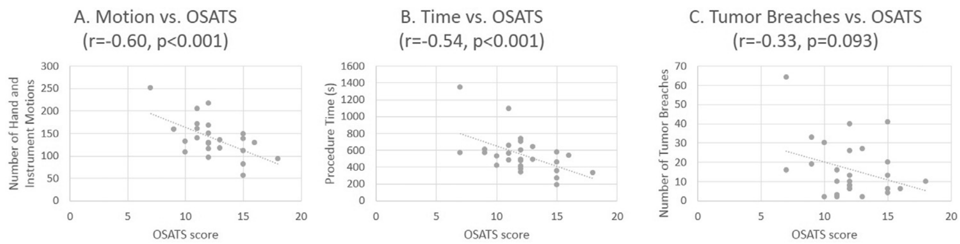

FIGURE 5. Correlation coefficients.

intermediate, and experienced groups, providing evidence of consequences for Surgery Tutor scores. Finally, we demonstrated evidence about relations to other variables via a “moderate” to “strong” correlation between motion and time metrics, and the gold-standard for assessment of technical proficiency—OSATS scores.

With the transition of postgraduate medical education toward a competency-based model, there is an increased need for reliable and objective assessment tools. Some of the current methods such as checklists and global rating scales have been well established in surgical education. The main advantages of procedure-specific checklists include objectivity, unambiguous expectations, and the opportunity to provide immediate and relevant feedback.[3] Checklists turn examiners into observers rather than interpreters of behavior, thereby making the assessment more objective and less subjective. Checklists show that the trainee knows the steps, but not if they are performing them well. The main advantages of global rating scales over checklists are their application to various procedures and the expert judgement of how well the task is performed.[3] The main disadvantages of global rating scales are the requirement to train the assessors in the use of the scale, introduction of some subjectivity during the assessment, and the difficulty of using these scales for formative feedback. In surgery, economy of motion is considered a key element in the progression toward proficiency.[28] Experts demonstrate greater motion economy than nonexperts,[29] and as trainees become more proficient, their motion economy becomes similar to the expert group.[25] Hand and instrument motion analysis systems, such as Imperial College Surgical Assessment Device, Advanced Dundee Psychomotor Tester, and Perk Tutor have demonstrated reliability and validity for objectively assessing trainees in performing percutaneous and laparoscopic procedures,[8,23] but this research has not been translated to open surgical procedures. Motion analysis systems also have the capability of providing immediate objective feedback which reduces the need for faculty presence. This can save faculty resources to be allocated to initial teaching and final assessments.[25]

Evidence supports that improved technical performance in the operating room results in improved patient outcomes.[30,31] This is especially important in oncologic surgery, where a positive margin can alter the survival outcomes.[32] Literature reports of positive margins rates for breast conserving surgery (BCS) using wire-guided localization vary widely, ranging from 12%-57%.[33] A positive margin rate of 29.1% was reported in a recent cohort study of 34,458 BCS cases.[34] These results are comparable to the 28% positive margin rate observed in our experienced group’s Surgery Tutor performance, which suggests that the fidelity of our phantom model is representative of real breast tissue for both palpable and nonpalpable tumors. We also reported a significantly lower positive margin rate for experienced surgeons (28%) compared to novice (68%) and intermediate level trainees (50%) ($p < 0.001$). We defined our experienced group as surgeons and senior residents that have performed more than 40 soft-tissue tumor resections in the preceding 5 years. This was based on studies of positive margin rates for BCS in the operating room that have reported a significantly lower rate of margin positivity after a surgeon performs 40 tumor resections.[21]

The OSATS scale is often used as the gold standard in assessing technical skills in the operative setting. We chose the OSATS scale as a comparator to show evidence of relations of Surgery Tutor scores to other variables. The OSATS scale has previously been demonstrated to be valid for assessment of technical proficiency. We used a single experienced reviewer who was blinded to the identity and group assignments of the participants in our study. There was a significant difference between OSATS scores for novice, intermediate, and experienced study participants. In our study, the novice, intermediate, and experienced groups scored on average 50%, 60%, and 72.5% respectively (Figure 4), which is similar to the OSATS data presented for PGY 1−6 residents by Regeher.[3]

Bias is created in our study by assigning participants to groups a priori based on self-reported case volumes. Although high case volumes have been associated with improved surgical outcomes,[35,36] volume alone does not equal proficiency. The OSATS scores correlated significantly with group assignment; however, there was overlap in scores between groups, including between novice and experienced. This reinforces the findings that volume alone does not provide evidence of proficiency and further justifies the need for additional assessment adjuncts including those that provide objective measurements.

In relation to other variables, Surgery Tutor demonstrated a "moderate" to "strong" correlation when hand/instrument motion and time were compared to the OSATS scores. When compared to number of tumor breaches, there was only a "weak" correlation with the OSATS scores. This is likely because the OSATS scale and number of tumor breaches measure different components of the construct of "technical proficiency." Thus, Surgery Tutor metrics would be extremely useful during training in principals of surgical oncology as we would be able to provide formative feedback about the number of tumor breaches to our trainees as they learn BCS in the laboratory and in the operating room.

There were limitations in our study related to the simulated phantom model. The plastisol material used to construct the phantoms does not adequately reflect the deformability of breast tissue, and the fragility of the "skin" layer of the phantom did not allow for appropriate flap creation and subcutaneous closure. This lack of realism made the assessment of "respect for tissue" in the OSATS scores challenging, as the video reviewer commented that this category was difficult assess. Due to these discrepancies, performance metrics generated by the Surgery Tutor cannot be generalized to the clinical setting without further testing.

Another limitation is the technology and resources required for Surgery Tutor, including motion sensors, EM field generator, control box, and laptop, which may limit widespread adoption of this assessment technique. While the motion-tracking platform does not require direct observation by an expert during the procedure, at the current time a technician is still required to set up the motion tracking equipment and oversee its use. Future potential developments in this area include an automated all-in-one system similar to the virtual reality endoscopy and laparoscopy trainers, so that trainees may use the Surgery Tutor platform and assessment technique independently in a simulation laboratory.

Surgery Tutor is useful as a method of assessing technical proficiency in an objective manner for soft-tissue resections. It has the potential to provide trainees with information regarding their economy of motion, number of tumor breaches, and duration of procedure, which can be used as feedback to guide further practice sessions and progression to the next training level. The metrics generated by Surgery Tutor also have the potential to be utilized as proficiency criteria. For example, as novices progress in their training, they could be expected to obtain intermediate and eventually expert level metrics. The purpose of Surgery Tutor is not as a replacement of faculty feedback and assessment, but as an adjunct for objective assessment and feedback during simulated practice sessions. This saves faculty resources for clinical teaching and final summative assessments. Future work will require validation of the transferability of technical proficiency demonstrated using Surgery Tutor in the simulated setting to the clinical environment.

CONCLUSION

We demonstrated construct validity of the scores obtained by Surgery Tutor for assessment of technical proficiency in soft-tissue tumor resections in a simulated setting. As a result, Surgery Tutor has the potential to be used for objective assessment and provision of formative feedback for surgical trainees learning open soft-tissue tumor excisions in a simulated setting.

APPENDIX

Table A

TABLE A. Observed Structured Assessment of Technical Skill (OSATS) Rating Scale. Anchors Taken From http://www.csats.com/osats/

	1	2	3	4	5
Respect for tissue	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissue appropriately with minimal damage
Time and motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument handling	Repeatedly makes tentative or awkward moves with instruments		Competent use of instruments although occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness
Flow of operation and forward planning	Frequently stopped operating or needed to discuss next move		Demonstrated ability for forward planning with steady progression of operative procedure		Obviously planned course of operation with effortless flow from one move to the next

REFERENCES

1. Harris KA, Frank JR. Competence by design: reshaping Canadian medical education. [Internet] Royal College of Physicians and Surgeons of Canada [cited 2018 July 18]. Available from: www.royalcollege.ca/rcsite/documents/educational-strategy-accreditation/royal-college-competency-by-design-ebook-e.pdf 2014.

2. Holmboe ES, Edgar L, Hamstra S. The milestones guidebook. [Internet] American Council of Graduate Medical Education [cited 2018 July 18]. Available from: https://www.acgme.org/Portals/0/Milestones-Guidebook.pdf 2016.

3. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med*. 1998;73:993–997.

4. Zevin B, Bonrath EM, Aggarwal R, Dedy NJ, Ahmed N, Grantcharov TP. Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg*. 2013;216:955–965.

5. Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med*. 1996;71:1363–1365.

6. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg*. 2005;190:107–113.

7. Beard JD. Objective assessment of technical surgical skills. *Br J Surg*. 2010;97:987–988.

8. Datta V, Mackay S, Mandalia M, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *J Am Coll Surg*. 2001;193:479–485.

9. Oropesa I, Sánchez-González P, Lamata P, et al. Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. *J Surg Res*. 2011;171:e81–e95.

10. Mason JD, Ansell J, Warren N, Torkington J. Is motion analysis a valid tool for assessing laparoscopic skill? *Surg Endosc*. 2013;27:1468–1477.

11. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med*. 2006;355:2664–2669.

12. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg*. 2008;247:372–379.

13. Ghaderi I, Manji F, Park YS, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. *Ann Surg*. 2015;261:251–262.

14. Van Der Vleuten C, Schuwirth L. Assessing professional competence: From methods to programmes. *Med Educ*. 2005;39:309–317.

15. Jakimowicz JJ, Buzink S. Training curriculum in minimal access surgery. In: Francis N, Fingerhut A, Bergamaschi R, Motson R, eds. Training in minimal access surgery, Editor 2015, Springer.

16. Zevin B, Dedy NJ, Bonrath EM, Grantcharov TP. Comprehensive simulation-enhanced training curriculum for an advanced minimally invasive procedure: a randomized controlled trial. *Surg Obes Relat Dis*. 2017;13:815–824.

17. Van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. Objective assessment of technical surgical skills. *Br J Surg*. 2010;97:972–987.

18. Brydges R, Kurahashi A, Brümmer V, Satterthwaite L, Classen R, Dubrowski A. Developing criteria for proficiency-based training of surgical technical skills using simulation: changes in performances as a function of training year. *J Am Coll Surg*. 2008;206:205–211.

19. Aggarwal R, Undre S, Moorthy K, Vincent C, Darzi A. The simulated operating theatre: comprehensive training for surgical teams. *BMJ Qual Saf*. 2004;13(suppl 1):i27–i32.

20. Palter VN, Orzech N, Reznick RK, Grantcharov TP. Validation of a structured training and assessment curriculum for technical skill acquisition in minimally invasive surgery: a randomized controlled trial. *Ann Surg*. 2013;257:224–230.

21. Toth D, Varga Z, Sebo E, Török M, Kovacs I. Predictive factors for positive margin and the surgical learning curve in non-palpable breast cancer after wire-guided localization—prospective study of 214 consecutive patients. *Pathol Oncol Res*. 2015;22:209–215.

22. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84:273–278.

23. Clinkard D, Holden M, Ungi T, et al. The development and validation of hand motion analysis to

evaluate competency in central line catheterization. *Acad Emerg Med*. 2015;22:212-218.

24. Ungi T, Sargent D, Moult E, et al. Perk Tutor: An open-source training platform for ultrasound-guided needle insertions. *IEEE T Bio-Med Eng*. 2012;59:3475-3481.

25. Yeo CT, Davison C, Ungi T, Holden M, Fichtinger G, McGraw R. Examination of learning trajectories for simulated lumbar puncture training using hand motion analysis. *Acad Emerg Med*. 2015;22:1187-1195.

26. Ungi T, Gauvin G, Lasso A, et al. Navigated breast tumor excision using electromagnetically tracked ultrasound and surgical instruments. *IEEE Trans Bio-Med Eng*. 2016;63:600-606.

27. Messick S. Standards of validity and the validity of standards in performance assessment. *Educ Meas-Issues Pract*. 1995;14:5-8.

28. Grober ED, Roberts M, Shin EJ, Mahdi M, Bacal V. Intraoperative assessment of technical skills on live patients using economy of hand motion: establishing learning curves of surgical competence. *Am J Surg*. 2010;199:81-85.

29. Datta V, Chang A, Mackay S, Darzi A. The relationship between motion analysis and surgical technical assessments. *Am J Surg*. 2002;84:70-73.

30. Birkmeyer JD, Finks JF, O'reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med*. 2013;369:1434-1442.

31. Fecso AB, Bhatti JA, Stotland PK, Quereshy FA, Grantcharov TP. Technical performance as a predictor of clinical outcomes in laparoscopic gastric cancer surgery. *Ann Surg*. 2018. [Epub ahead of print].

32. Meric F, Mirza NQ, Vlastos G, et al. Positive surgical margins and ipsilateral breast tumor recurrence predict disease−specific survival after breast−conserving therapy. *Cancer*. 2013;97:926-933.

33. Chan BKY, Wiseberg-Firtell JA, Jois RHS, Jensen K, Audisio RA. Localization techniques for guided surgical excision of non-palpable breast lesions. *Cochrane Database Syst Rev*. 2015;12:CD009206. https://doi.org/10.1002/14651858.CD009206.pub2.

34. Van Leeuwen MT, Falster MO, Vajdic CM, et al. Reoperation after breast-conserving surgery for cancer in Australia: statewide cohort study of linked hospital data. *BMJ Open*. 2018;8:e020858. https://doi.org/10.1136/bmjopen-2017-020858.

35. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med*. 2002;346:1128-1137.

36. Papachristofi O, Jenkins D, Sharples LD. Assessment of learning curves in complex surgical interventions: a consecutive case-series study. *Trials*. 2016;17:266.

SUPPLEMENTARY INFORMATION

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jsurg.2018.11.005.