Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Beyond task time: automated measurement augments fundamentals of laparoscopic skills methodology

Timothy M. Kowalewski, PhD,[a,*] *Lee W. White, PhD,*[b] *Thomas S. Lendvay, MD, FACS,*[c] *Iris S. Jiang, BS,*[b] *Robert Sweet, MD, FACS,*[d] *Andrew Wright, MD,*[e] *Blake Hannaford, PhD,*[f] *and Mika N. Sinanan, MD, PhD, FACS*[e]

[a] *University of Minnesota Mechanical Engineering, Minneapolis, MN*
[b] *University of Washington Bioengineering, Seattle, WA*
[c] *University of Washington Urology and Seattle Children's Hospital, Seattle, WA*
[d] *University of Minnesota Urology, Minneapolis, MN*
[e] *University of Washington Surgery, Seattle, WA*
[f] *University of Washington Electrical Engineering, Seattle, WA*

ARTICLE INFO

Article history:
Received 17 February 2014
Received in revised form
12 May 2014
Accepted 27 May 2014
Available online 4 June 2014

Keywords:
Laparoscopy
Simulation
Education
Objective metrics
FLS
Validation
Grasping force

ABSTRACT

Background: Laparoscopic psychomotor skills are challenging to learn and objectively evaluate. The Fundamentals of Laparoscopic Skills (FLS) program provides a popular, inexpensive, widely-studied, and reported method for evaluating basic laparoscopic skills. With an emphasis on training safety before efficiency, we present data that explore the metrics in the FLS curriculum.
Materials and methods: A multi-institutional ($n = 3$) cross-sectional study enrolled subjects ($n = 98$) of all laparoscopic skill levels to perform FLS tasks in an instrumented box trainer. Recorded task videos were postevaluated by faculty reviewers ($n = 2$) blinded to subject identity using a modified Objective Structured Assessment of Technical Skills (OSATS) protocol. FLS scores were computed for each completed task and compared with demographically established skill levels (training level and number of procedures), video review scoring, and objective performance metrics including path length, economy of motion, and peak grasping force.
Results: Three criteria used to determine expert skill, training and experience level, blinded review of performance by faculty via OSATS, and FLS scores, disagree in establishing concurrent validity for determining "true experts" in FLS tasks. FLS-scoring exhibited near-perfect correlation with task time for all three tasks (Pearson $r = 0.99$, 1.00, 1.00 with $P < 0.00000001$). FLS error penalties had negligible effect on FLS scores. Peak grasping force did not correlate with task time or FLS scores.
Conclusions: FLS technical skills scores presented negligible benefit beyond the measurement of task time. FLS scoring is weighted more toward speed than precision and may not significantly address poor tissue handling skills, especially regarding excessive grasping force. Categories of experience or training level may not form a suitable basis for establishing proficiency thresholds or for construct validity studies for technical skills.

* *Corresponding author*. University of Minnesota, Minneapolis, MN 55455. Tel.: +1 612 626 0054; fax: +1 612 625 4344.
E-mail address: tmk@uw.edu (T.M. Kowalewski).
0022-4804/$ – see front matter
http://dx.doi.org/10.1016/j.jss.2014.05.077

1. Introduction

The Accreditation Council of Graduate Medical Education specifically identifies "technical competence in conducting surgical procedures" as a component in its core competencies under patient care and practice-based learning [1]. Professional surgical organizations responsible for certification of technical knowledge and proficiency and/or competency also need means for accurately assessing surgical skills [2]. Skill assessment systems such as Objective Structured Assessment of Technical Skills (OSATS) or the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills and its evolution into the Fundamentals of Laparoscopic Skills (FLS) have provided a framework to analyze surgical skill [3–6]. In 2004, the FLS assessment certification process was launched and after 5 y of its existence, almost 3000 clinicians have become certified [7]. We seek to build on this success by evaluating the metrics within existing FLS scoring and defining options for possible improvement.

The psychomotor skills aspect of FLS measures competency in the domain of laparoscopic surgical skills with several different measures. The FLS score for each task is derived from a formula based on task time and numbers of errors (Table 1), and is weighted such that time is heavily weighted and usually a dominant factor in the formula [8]. One of the limitations of the technical skills portion of FLS is that there are few objective metrics. A heavy focus on task time may mask other important measures of safety and proficiency. Task time alone is unlikely to accurately reflect the quality of the surgical performance and rushing to complete a task in the fastest possible time may negatively impact quality. Meanwhile, tissue handling as measured by grasping force on tissue as a metric is not measured. It directly correlates with tissue damage. This can lead to such poor clinical correlates such as a tear in a critical anatomic structure, a leaking anastomosis, or a physiologically significant stricture.

Although the discriminating power of FLS clearly separates experienced and inexperienced subjects [9–12], more granular measures of surgical performance such as economy of motion (EoM), grasp forces, and tool motion characteristics may also be important independent measure of performance but have not been evaluated. Physical model-based laparoscopic box trainers are not capable of capturing or reporting these metrics routinely. Moreover, there is no rigorous and repeatable method established to carry out identical performance measures on different platforms [1,13–16].

To address these issues, we created a laparoscopic box trainer that uses real tools and physical models in dry or wet laboratory tasks but also contains instrumentation for automated capture and analysis of tool motion, EoM, and grasping forces [17]. Building on prototype models (the red [18], and blue dragon [19]) this instrumented trainer box is a scaled-down, table-top version of the larger scale platforms, which were successfully used as a surgical skills research tool in live porcine surgery. De *et al.* [20] used a related grasper mechanism to establish peak grasp force thresholds that result in tissue damage. The red dragon prototype, along with the established hidden Markov model-based scoring methodology [21] as licensed and commercialized by Simulab Corporation (Seattle, WA), was incorporated into the EDGE (Fig. 1). To our knowledge, EDGE is the only dry laboratory reality-based box trainer that can obtain high-accuracy tool motion (position and orientation) measurements along with grasping force and synchronized video [22]. We used EDGE as a research platform to evaluate different measure of quantitative skill in FLS tasks.

Table 1 – Equations used to compute FLS scores [23,24,28].

FLS task	FLS score
PegTx	$FLS_{Peg} = (300 - t - 17E_{dr})/237$
Cutting	$FLS_{Cut} = (300 - t - 2E_a)/280$
Suturing	$FLS_{Sut} = (600 - t - E_{pd} - E_g - E_q)/520$

The goal of this work is to evaluate the training goals and scoring methods of FLS psychomotor skill scoring and compare them with a broader set of performance criteria available in EDGE. We hypothesize that existing FLS scoring with its emphasis on task time may unintentionally promote speed over careful tissue handling and lead to less reliable or less accurate global measures of skill. We further propose that

Fig. 1 – The EDGE Platform was developed by Simulab Corporation and is based on a mechanism developed by the University of Washington Biorobotics laboratory. It consists of a pair of interchangeable surgical tools whose motion is constrained to rotate about a fulcrum the same way laparoscopic instruments are constrained by their access ports. (Color version of figure is available online.)