JMIG The Journal of Minimally Invasive Gynecology

Review Article

Gynecologic Endoscopy Skills Training and Assessment: Review

Rasiah Bharathan, MBBS, MSc, MRCS, MRCOG*, Thomas Setchell, MRCOG, Tariq Miskry, MRCOG, Ara Darzi, FRCS, and Rajesh Aggarwal, FRCS, PhD

From the Department of Surgery and Cancer, Imperial College (Drs.Bharathan, Darzi, and Aggarwal), Department of Obstetrics and Gynecology, Women's Centre, Imperial College Healthcare NHS Trust (Drs.Setchell and Miskry), St. Mary's Hospital, London, Nuffield Department of Obstetrics and Gynecology, Oxford University Hospitals NHS Trust, Oxford (Dr.Bharathan), UK, and Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Dr.Aggarwal).

ABSTRACT Training in and assessment of endoscopic skills is currently undergoing a period of evolution. Several recognized factors driving this evolution include working pattern, training opportunities, cost, and patient safety. In addition, the need to continuously monitor competence is punctuated by the rapid technologic changes and rising consumer expectation. These challenges present an opportunity to positively enhance the learning and performance of surgical practice. Journal of Minimally Invasive Gynecology (2014) 21, 28–43 © 2014 AAGL. All rights reserved.

Keywords: Assessment; Curriculum; Endoscopy skills; Gynecological surgery; Training

DISCUSS You can discuss this article with its authors and with other AAGL members at **http://www.AAGL.org/jmig-21-2-JMIG-D-13-00225R1**

Use your Smartphone to scan this QR code and connect to the discussion forum for this article now*

* Download a free QR Code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

The Halstedian model has guided traditional surgical training for more than a hundred years [1]. The model relies on the trainee developing the skills and acquiring knowledge over protracted periods, with exposure dependent on chance encounter of pathologic conditions, trainer willingness, and availability of operating room resources. The operating room is not the ideal educational environment because it is not always feasible to provide timely objective feedback [2]. Furthermore, surgery is a time-critical activity, and this is compounded in that teaching *per se* can be perceived as stressful [3]. With increasing nonsurgical treatments (e.g., ectopic pregnancyand menorrhagia) and shortened training period, the efficacy of the traditional training method has been eroded. Safety and cost-efficiency initiatives have driven these changes.

Although operative laparoscopy in gynecology was popularized in the 1970s, in particular with sterilization, the prevailing interest in modern surgical education was brought into focus in the early 1990s with the introduction of laparoscopic surgery in the main stream. The Hospital Association of New York was one of the first organizations to recognize the need to review advanced laparoscopic procedures before hospital credentialing [4]. The need for structured training and credentialing was recognized more than 20 years ago [5], with simultaneous developments in North America [6], Europe [7], and Australia [8]. These developments prompted a paradigm shift to move the learning curve of endoscopy away from the operating room into simulation laboratories.

In a recent editorial review, synthetic simulators and the value of organic simulators in laparoscopic and hysteroscopic surgery have been described [9]. In that review, surgical simulation is supported by outlining the concept of simulator validation and its relevance to simulation-based training [9]. Another article has provided a brief overview of laparoscopic simulation and its relevance to curriculum design [10].

Material and Methods

The present article, which is not intended as a systematic review of any single aspect of simulation-based education,

Dr. Aggarwal is supported by a Clinician Scientist award from the UK National Institute for Health Research.

Corresponding author: Rasiah Bharathan, MBBS, MSc, MRCS, MRCOG, Department of Surgery and Cancer, Imperial College, 10th Floor, QEQM Hospital, Paddington, London W2 1NY, UK.
E-mail: r.bharathan@imperial.ac.uk

Submitted April 10, 2013. Accepted for publication July 30, 2013.
Available at www.sciencedirect.com and www.jmig.org

1553-4650/$ - see front matter
http://dx.doi.org/10.1016/j.jmig.2013.07.016

provides a comprehensive overview of gynecologic endoscopic simulators, both laparoscopic and hysteroscopic, including the various techniques of synthetic models and of animal and cadaver models. This is complemented by a discussion of the methods of skill assessment. In addition the key concepts of learning curve, simulator efficacy, skill transference, and curriculum design are discussed.

A literature search was performed using both Medical Subject Headings (MeSH) and keywords; the search strategy is given in the Appendix. Medline and EMBASE databases were searched frominception to May 2013; further references were obtained via hand searching of the bibliography and selected review articles.

Simulators

Introduction of simulators first requires demonstration of the subjective entities of fidelity and content validity. Construct, concurrent, and predictive validities are objective dimensions that determine the efficacy of a simulator within a curriculum (Table 1). The range of simulators includes synthetic, virtual reality (VR), cadavers and animal models [11].

Synthetic Simulators

Laparoscopy Simulators

Synthetic models include box trainers and video trainers, and they have the advantage of low cost and portability (Fig. 1). Although clinical outcome would provide the best metrics of performance, in the context of simulation-based education, alternative metrics are appropriate. Validation of laparoscopic simulators is given in Table 2.

Several custom-built and commercially available synthetic simulators have been used to assess validity of basic and procedural tasks [12,17,18,20,24,26–29]. Content validity is only implied in some studies because they were designed or chosen by the faculty to reflect the skills required of trainees. Most of the studies have recruited experts (i.e., attending physicians or faculty recognized as laparoscopists) [17,20,24,26,27,30], and a few have attempted to ascertain construct validity among 3 categories of participants based on designation [24,26,28,30]. In 2 studies performed by Kolkman et al [17,27], the simulator and the 5 basic tasks were identical. In the first study, 8 experts and 111 residents were recruited [27], and in their second study, 5 experts and 18 students participated [17]. Findings of both studies concurred that a composite score comprising time and error scores was indeed a construct valid parameter. Secondary finding by Kolkman and colleagues were that previous laparoscopic experience, but not previous simulator experience, was correlated with performance [27].

In the largest validation study to date, Molinas et al [20] tested 3 basic tasks: camera navigation, camera navigation and forceps handling, and forceps handling and bimanual coordination. A pilot study demonstrated construct validity (time and timed bead transfer) between 10 experts and 14 novices; this finding was confirmed in a larger study comprising 42 experts and 241 novices from 34 countries. Speed of task completion does not equate with quality of performance in the clinical setting; however, time and timed activities have proved to be the most common construct valid parameter in assessing basic [17,20,24,26–30] and procedural [18] tasks. In addition, error scores, when combined with completion time, have demonstrated composite score as a construct valid parameter [17,24,27,29]. Observational objective assessment tools are now widely applied across specialties [31].

Table 1

Definitions of validity and reliability

Dimension	Definition
Face validity (or fidelity)	Assessment undertaken by experts to determine whether the device or tool measures what it purports to perform. This is a subjective dimension.
Content validity	Systematic review of the constituent parts of the device or tool to determine whether the components are appropriate insofar as overall cohesion of the tool. This is a judgmental process conducted by experts.
Construct validity	Extent to which the device or tool can measure the quality and/or the quantity it is designed to measure. In the context of surgical simulation, an application would differentiate a cohort in accordance with level of experience.
Concurrent validity	Extent to which the assessment outcome concurs with that of another device or tool assessing the same construct.
Predictive validity	Ability of the device or tool to predict performance in the actual performance of the task, i.e., operative performance in a human patient.
Retest reliability	Refers to the stability of the outcome generated on multiple assessment points separated in time. A reliability coefficient of 0.8 is considered acceptable; that between 0.5 and 0.8 is moderately useful; correlation coefficient <0.5 demonstrates poor stability.
Interrater reliability	In the context of observational assessment, the outcome of a process conducted by at least 2 separate assessors must demonstrate a level of agreement. The reliability coefficient is as for retest reliability.
Internal consistency	Degree of correlation between the various metrics of a device or tool.

Adapted from Sugden and Aggarwal [11].