

# Construct Validity of Fresh Frozen Human Cadaver as a Training Model in Minimal Access Surgery

Mitesh Sharma, MS, David Macafee, MA, MD, Nagarajan Pranesh, MD, Alan F. Horgan, MD

## ABSTRACT

**Background:** The construct validity of fresh human cadaver as a training tool has not been established previously. The aims of this study were to investigate the construct validity of fresh frozen human cadaver as a method of training in minimal access surgery and determine if novices can be rapidly trained using this model to a safe level of performance.

**Methods:** Junior surgical trainees, novices (<3 laparoscopic procedure performed) in laparoscopic surgery, performed 10 repetitions of a set of structured laparoscopic tasks on fresh frozen cadavers. Expert laparoscopists (>100 laparoscopic procedures) performed 3 repetitions of identical tasks. Performances were scored using a validated, objective Global Operative Assessment of Laparoscopic Skills scale. Scores for 3 consecutive repetitions were compared between experts and novices to determine construct validity. Furthermore, to determine if the novices reached a safe level, a trimmed mean of the experts score was used to define a benchmark.

Mann-Whitney *U* test was used for construct validity analysis and 1-sample *t* test to compare performances of the novice group with the benchmark safe score.

**Results:** Ten novices and 2 experts were recruited. Four out of 5 tasks (nondominant to dominant hand transfer; simulated appendectomy; intracorporeal and extracorporeal knot tying) showed construct validity. Novices'

scores became comparable to benchmark scores between the eighth and tenth repetition.

**Conclusion:** Minimal access surgical training using fresh frozen human cadavers appears to have construct validity. The laparoscopic skills of novices can be accelerated through to a safe level within 8 to 10 repetitions.

**Key Words:** Cadaver training, Laparoscopic skills training, Construct validity.

## INTRODUCTION

The history of using human cadavers for learning anatomy can be traced back to 16th century AD when Andrea Vesalius, referred to as the Father of Anatomy, published an epic on anatomical examination in 1543 AD, titled *de humani corporis fabrica* (on the fabric of the human body).<sup>1</sup> Until recently, cadavers have been used for dissection and demonstration of anatomy but not to train in surgical procedures.<sup>1</sup> This concept took a major revamp when the United Kingdom (UK) Human Tissue Act 2004 established standards and provided guidance to clinicians to carry out education and training using cadaveric materials. This allowed for research, education, and surgical skills training to be carried out on donors who bequeathed their bodies to medical science.

Prior to implementation of any training tool, evaluation and validation of the tool and its parameters are mandatory.<sup>2</sup> Construct validity can be defined as "evaluating a testing instrument based on the degree to which the test item identifies the quality, ability, or trait it was designed to measure."<sup>3</sup> This is usually done by measuring the performance in 2 groups that are hypothesized to differ in the skill being measured by the instrument (e.g., experienced surgeons and novices).<sup>2,4,5</sup> There is a paucity of studies on human cadavers as a method of simulated training for laparoscopic procedures and none dealing specifically with construct validity or evaluating "fresh cadaver" as a training tool.<sup>6-9</sup>

There are multiple factors that will aid rapid acquisition of technical surgical skills, but evidence of the construct

Newcastle Surgical Training Centre, Department of General Surgery, Freeman Hospital NHS Trust, Newcastle Upon Tyne, UK (Drs Sharma, Pranesh, Horgan).

Department of General Surgery, The James Cook University Hospital, Middlesbrough, UK (Mr. Macafee).

The authors would like to thank Mr. Tony Fouweather, Statistician, Newcastle University, for his help with statistical analysis. The authors also thank Dr. Roger Searle, Director, Anatomy and Clinical Skills, Newcastle University, and Mr Julian Hance, SpR, Leeds Royal Infirmary, for their invaluable advice.

Address correspondence to: Mitesh Sharma, Research and Education Fellow, Newcastle Surgical Training Centre, Department of General Surgery, Freeman Hospital NHS Trust, Newcastle Upon Tyne, NE7 7DN, UK. Telephone: 0044-7515113584; E-mail: dr\_miteshsharma@yahoo.co.uk

DOI: 10.4293/108680812X13462882735818

© 2012 by JSLS, Journal of the Society of Laparoendoscopic Surgeons. Published by the Society of Laparoendoscopic Surgeons, Inc.

validity of fresh frozen cadaver (FFC) use in operative skills training has not been established to date. With both access to and availability of human cadavers being strictly controlled, it is also important to know how much exposure (e.g., number of repetitions of a particular procedure) is needed to reach an acceptable level of safety to operate on live subjects.

The primary aim of the study was to assess the construct validity of FFC as a method of minimal access surgical skills training. A secondary aim was to identify how much exposure to this training medium, in terms of number of repetitions of a task, were required to see the skills of the novice rise to a predetermined safe level.

## METHODS

This was a single-center trial conducted in a Human Tissue Act-approved fresh cadaver laboratory in a UK teaching hospital. Appropriate ethical approval was obtained from the local research ethics committee. Recruitment to the study was through global E-mails sent to surgical trainees. Those who had previously performed >3 laparoscopic procedures were excluded from the novice group.

Ten junior surgical trainees (<2 y of surgical experience) who were novices in laparoscopic surgery and 2 experts, fully trained general surgeons (postlaparoscopic surgery fellowship and >100 laparoscopic procedures) with their practice composed primarily of minimal access surgery, were recruited. Neither experts nor novices had practiced on the cadaver model or were familiar with the structure of our curriculum before the study commenced.

### Cadaver-Based Technical Skills Exercise

A set of 5 tasks were improvised from the Fundamentals of Laparoscopic Surgery (FLS) curriculum, a valid and reliable curriculum developed by the Society of American Gastrointestinal and Endoscopic Surgeons.<sup>10</sup>

Task 1. Nondominant to dominant hand sharp peg transfer: Colored peg, with sharp end, to be lifted from 1 paracolic gutter using nondominant hand; transferred in midair to dominant hand and placed in other paracolic gutter

Task 2. Dominant to nondominant hand sharp peg transfer: Above-mentioned exercise to be started with dominant hand

Task 3. Simulated appendectomy: Mesentery to be cut flush off 5-cm to 6-cm segment of divided small bowel

(avoiding any perforations to bowel). This segment of bowel to be endolooped and divided. This task incorporates pattern cut and endoloop tasks of FLS

Task 4. Intracorporeal knot tying: The mesenteric rifts to be closed using 2-Vicryl suture with 3 square knots

Task 5. Extracorporeal knot tying: The mesenteric rifts to be closed using 2-PDS suture with 4 square knots

## Cadaver Practice and Assessments

### Cognitive Learning

Novices and experts watched a live demonstration of the above-mentioned 5 tasks, performed by a national laparoscopic trainer, followed by an interactive session to clarify any doubts. In addition, a DVD of this demonstration was available for the trainees and experts to refer to during the technical skills exercise sessions, if needed.

### Technical Skills Exercise

Novices performed 10 repetitions of 1 task before they progressed on to perform the rest of the tasks mentioned above.

Participants worked in pairs (as operator and camera assistant). The operator took a rest for 2.5 min after each repetition. In addition, the operator and camera assistant took a break for 20 min and swapped their roles after completing 10 consecutive repetitions of 1 task or after 1 h of practice session block, whichever was earlier. This was to ensure adequate rest and allow distributed training.<sup>11,12</sup> The whole exercise was spread over 2 d for trainees.

Experts performed 3 repetitions of each task in the similar manner. An invigilator was present during training sessions to ensure no undue help was taken from the assistant and to document the number of times any participant referred to the demonstration DVD. All repetitions were video recorded for later analysis by 2 blinded assessors.

### Assessments

All of the experts' and trainees' recorded procedures were split into individual repetition clips using iMovies (Apple, USA) software. These were then randomised and coded using a computer-generated random number list to ensure blinding of assessors. The expert and trainee videos were kept in the same pool of clips. Hence, the assessor (postlaparoscopic fellowship consultant surgeon) did not know the identity of either the performer or the order of repetition.

A modified version of validated Global Operative Assessment of Laparoscopic Skills (GOALS) scale was used to score the performance.<sup>13–15</sup>

Autonomy was to be assessed as per the domain in the GOALS scale, and this was to be achieved by assessing the help required or requested from the invigilator or from the need to reference the training DVD. This was to be mentioned on coded video clips for the assessors to make an inference. As no novice referred to the DVD nor asked for help from the invigilator, the autonomy domain was removed from the assessment.

To check the reliability of assessments, scores of the principal assessor were compared with a second assessor for the task of intracorporeal knot tying. The second assessor was also equivalent to the primary assessor in terms of surgical experience (postlaparoscopic fellowship consultant grade surgeon) and scored these blinded clips using same assessment scale.

### Proficiency Level Development

The benchmark score for each task was obtained by calculating the mean performance score of both experts for the 3 repetitions, excluding the values lying beyond 2 standard deviations from the mean. The trimmed mean was used as the benchmark score for safe performance.

### Performance Evaluations

Any significant difference in consecutive scores for the first 3 repetitions between and expert groups was taken as a measure of construct validity.

Comparison of each of the 10 repetition scores of the novices' group with the benchmark score was used to assess the number of task-specific repetitions required for trainees to reach the benchmarked "safe level" on all construct valid tasks.

### Statistical Analysis

The data were analysed with SPSS version 17 (SPSS, Chicago, Illinois, USA). The Mann-Whitney *U* test was used to compare the performance of novices and experts to determine construct validity. A 1-sample *t* test was used to compare benchmark score with the novice group's scores for 10 repetitions. Inter-rater reliability was assessed using Kendall's  $\tau$ -b and Spearman's rho tests.  $P < .050$  was considered statistically significant.

## RESULTS

### Construct Validity

The median (IQR) performance scores of experts was significantly higher than those of novices for nondominant to dominant hand peg transfer [15(13) vs 12(2);  $P = .006$ ]; simulated appendicectomy [18(3) vs 14(3);  $P = .005$ ]; intracorporeal knot tying [15.5(5) vs 12(3);  $P = .001$ ]; and extracorporeal knot tying [16.5(3.75) vs 14(3.25);  $P = .033$ ].

The median (IQR) scores for the task of dominant to nondominant hand peg transfer showed no difference between the groups [11.5(2) vs 11(4);  $P = .068$ ] (**Figures 1A to 1E**).

### Calculation of Benchmark Score Using Modified GOALS Scale (Max Score = 20)

The scores achieved by experts for each task are shown in **Table 1**. The mean expert scores for the tasks of simulated appendicectomy, intracorporeal and extracorporeal knots were readjusted by trimming the values lying outside 2 SD of the group mean to obtain the benchmark scores.

### Comparison of Novices' Scores with Benchmark Score

The performance score of the novice group did not show any significant difference for the task of dominant to nondominant hand peg transfer compared to the benchmark score even for the first repetition [mean novices' performance score 10.7, benchmark score 11.25;  $P = .569$ ] or for subsequent repetitions ( $P > .05$ ).

The mean scores and comparisons to benchmark scores for the remaining 4 tasks are shown in **Tables 2A to 2D**. The mean scores for the these tasks performed by novices became nonsignificant (or comparable to the benchmark scores) between the eighth to tenth repetitions for each task.

### Interrater Reliability

The association between ranks by 2 assessors was significant ( $P < .001$ ), indicating that the 2 assessors ranked the individual clips in a similar way (**Table 3**).

## DISCUSSION

Laparoscopic surgery training in the operating theater is often unstructured.<sup>16</sup> The present era of time-restricted

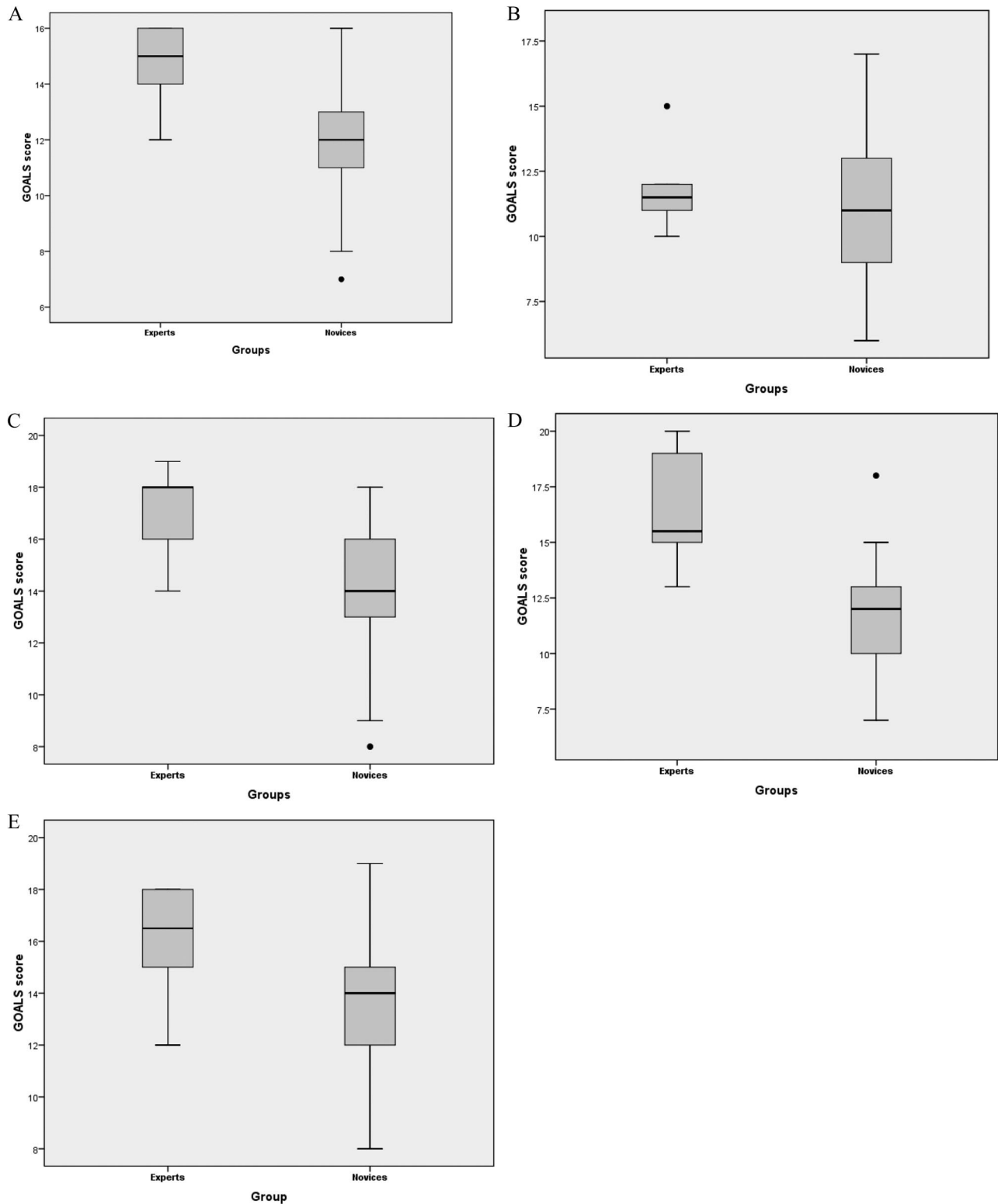


Fig. 1. Box and whisker plots for GOALS scores of novices (n=10) and experts (n=2) for 5 tasks. Boxes (interquartile range); bars (medians); whiskers (range excluding outliers); dots (outliers).

**Table 1.**  
Calculation of Bench Mark Score Using Modified GOALS Scale (max score=20)

Tasks	Expert 1 scores			Expert 2 scores			Expert Group's Mean	Standard Deviation (×2)	Trimmed Mean
	Rep. 1	Rep. 2	Rep. 3	Rep. 1	Rep. 2	Rep. 3			
Nondominant to dominant hand peg transfer	14	16	14	16	16	12	14.66	3.26	14.66
Simulated appendectomy	18	14	18	19	18	16	17.16	1.84	17.83
Intracorporeal knot tying	15	20	16	15	19	13	16.33	2.51	15.25
Extracorporeal knot tying	12	16	15	17	18	18	16	3.84	16.58

**Table 2A.**  
Nondominant to Dominant Hand<sup>a</sup>

Repetition	Mean	Mean Difference	95% CI of the Difference		P Value <sup>b</sup>
			Lower	Upper	
1	10.90	-3.760	-5.681	-1.839	.002
5	12.40	-2.260	-3.698	-0.821	.006
8	13.00	-1.660	-3.000	-0.311	.021
9	13.50	-1.160	-2.340	0.020	.053
10	14.50	-0.160	-1.9204	1.600	.842

<sup>a</sup>Test Value (Benchmark Score) = 14.66.

<sup>b</sup>P Value became nonsignificant on ninth repetition.

**Table 2B.**  
Appendectomy

Repetition	Mean	Mean Difference	95% CI of the Difference		P Value <sup>b</sup>
			Lower	Upper	
1	14.30	-3.530	-5.251	-1.808	.001
5	15.10	-2.730	-4.363	-1.097	.004
8	15.70	-2.130	-3.438	-0.821	.005
9	16.70	-1.130	-2.522	0.262	.100
10	16.20	-1.630	-3.375	0.115	.064

<sup>a</sup>Test Value (Benchmark Score) = 17.83.

<sup>b</sup>P value became nonsignificant on ninth repetition.

**Table 2C.**  
Intracorporeal Knot

Repetition	Mean	Mean Difference	95% CI of the Difference		P value
			Lower	Upper	
1	12.00	-3.250	-5.070	-1.430	.003
5	12.50	-2.750	-4.570	-0.930	.008
8	13.22	-2.028	-3.510	-0.550	.013
9	12.78	-2.472	-3.950	-0.990	.005
10	14.33	-0.917	-1.860	0.020	.055

<sup>a</sup>Test Value (benchmark score) = 15.25.

<sup>b</sup>P value became nonsignificant on tenth repetition.

**Table 2D.**  
Extra-corporeal Knot

Repetition	Mean	Mean Difference	95% CI of the Difference		P Value
			Lower	Upper	
1	13.50	-3.080	-5.140	-1.020	.008
5	13.30	-3.280	-5.250	-1.310	.004
8	15.50	-1.080	-3.050	0.890	.247
9	15.70	0.880	-2.230	0.470	.175
10	15.30	-1.280	-2.970	0.410	.120

<sup>a</sup>test value (benchmark score) = 16.58.

<sup>b</sup>P value became nonsignificant on eighth repetition.

training, requires trainees to acquire basic skills rapidly to progress to advanced training as soon as possible. Teaching complex skills like these on patients presents formidable obstacles as the teaching surgeon's role as an

assistant is more limited than in open surgery.<sup>7</sup> It has also been demonstrated that operating room training of junior surgeons is time consuming, resulting in increased cost.<sup>17,18</sup>

**Table 3.**  
Interrater Reliability Test for Video Clips Assessments

			Second Assessor	Principal Assessor
Kendall's tau-b	Second assessor	Correlation Coefficient	1.000	.520 <sup>a</sup>
		Sig. (2-tailed)		.000
		N	106	106
	Principal assessor	Correlation Coefficient	.520 <sup>a</sup>	1.000
		Sig. (2-tailed)	.000	
		N	106	106
Spearman's rho	Second assessor	Correlation Coefficient	1.000	.636 <sup>a</sup>
		Sig. (2-tailed)		.000
		N	106	106
	Principal assessor	Correlation Coefficient	.636 <sup>a</sup>	1.000
		Sig. (2-tailed)	.000	
		N	106	106

<sup>a</sup>Correlation is significant at the 0.01 level (2-tailed).

The value of practice on cadavers before proceeding to supervised live patient surgery and ultimately independent practice has been acknowledged by trainers.<sup>1</sup> Traditionally, embalmed human cadaver specimens have been used in medical school dissection classes. The stiffening and discoloration caused by full embalming are unavoidable but unacceptable for surgical training.<sup>19</sup> Trainees learning the visual and tactile components should work on tissues that have the same appearance and handling qualities as living tissue.<sup>19</sup> The closest anatomical model to resemble a live patient is fresh human cadaver. These issues make the perfect ground for training models like human cadavers to be used in training of surgeons to a safe level before embarking on to patients.

To provide a realistic learning experience, our lab uses human FFCs. FFC can be kept indefinitely in freezers with the temperatures between  $-17^{\circ}\text{C}$  and  $-20^{\circ}\text{C}$ . In order to defrost the cadavers for use, they are placed into a refrigerated area that is at a temperature between  $+3^{\circ}\text{C}$  and  $+5^{\circ}\text{C}$ . This takes up to 2 wk for the cadavers to defrost until they are ready for use. When this technique is used, the skeletal and visceral tissues retain the colors found in the living body. The compliance of tissues and the ease or difficulty in separating one structure from another resembles those of living tissue.

The usefulness of fresh cadavers to practice surgical skills is unique. The trainee has to practice maneuvers with real-life challenges, like proximity to other viscera and the inherent danger of injury to them; similar tissue compli-

ance; similar depth perception and abdominal pressures; and similar tactile feedback of tissues along with the challenge of dealing with a similar color scheme intraabdominally. All these cannot be replicated in any physical and virtual models of training.

A previous study on lightly embalmed cadavers for use in gynecology skills training evaluated the reduction in handling and manipulation times during procedure-specific maneuvers.<sup>7</sup> We measured the performance score rather than the speed of the operation, because it was considered more important for a junior trainee to perform a procedure following appropriate steps than simply encouraging speed of completion, which would be best learned during the next level of training. Efficiency of movements is certainly to be encouraged, the ultimate outcome of which would be more rapid completion of tasks in subsequent repetitions.

Eversbusch et al.<sup>20</sup> in their randomized trial on psychomotor training on virtual reality simulated colonoscopy, found that learning curves reached a plateau for experienced surgeons, senior trainees, and novices after the second, fifth, and seventh repetitions, respectively. Greater realism in training was shown for FFC than for virtual reality simulators in a previous study.<sup>21</sup> Hence, we made novices perform 10 repetitions of psychomotor skills on FFC, anticipating improvement within this period.

The GOALS scale, developed by McGill University Health Centre, Canada, can be used for evaluating performance

(with construct validity and reliability) on synthetic simulation tasks and live animal models.<sup>14</sup>

Gumbs et al.<sup>13</sup> concluded in their study that the same scale can be used to assess laparoscopic appendectomy with construct validity, and Vassiliou et al.<sup>15</sup> concluded that it can be used to assess laparoscopic skills based on videotaped performance. The GOALS scale modification (removal of autonomy) was necessary in our study to ensure blinding, because the performance was assessed by video tapes rather than direct observation. Autonomy is an important domain to assess performance of trainee, and a provision was made during our experiment to record the level of autonomy for trainees. This was captured by a record kept by the invigilator, recording the number of times the trainee referred to the demonstration DVD. This was to be mentioned on the coded video clip for the assessor to make an inference. Since no trainee or expert referred to the demonstration DVD, this domain was removed from the scale in our study. This also indicates that adequate cognitive training is of paramount importance to learn technical skills more effectively and should precede technical skills training. This could explain why our participants did not lack in autonomy. We recommend that if the assessments are done in a face-to-face setting, the domain of autonomy should be retained in the assessment scale.

Our method of establishing construct validity is consistent with previous studies on the subject.<sup>2,16,22</sup> The method of establishing a benchmark score using a trimmed mean score of 2 experts is also consistent with the previous study by Ritter et al.<sup>10</sup> who have established benchmark scores for FLS curriculum using 2 experts. Similarly, in order to help novices make the proficiency levels achievable in a reasonable amount of time and to compensate for the fact that the expert group consisted of only 2 experts, levels for all the tasks were set at 2 SD from the determined means. This allows for an increased chance of including the true mean performance scores of a larger sampling of expert surgeons while not setting the performance bar too high for the novices.<sup>10</sup>

The task of “dominant to nondominant hand peg transfer” did not show significant improvement with subsequent repetitions. This may be because it was too short a task to be effectively scored using an explicit GOALS scale. The other reason could be that the novices had performed 10 repetitions of nondominant to dominant hand peg transfer before embarking on the task. The similarity of the 2 tasks and repeated practice of that task could have resulted in improved dexterity skills of novices, sufficient enough to

negate any statistical differences compared to the expert group. The transfer of peg from nondominant to dominant hand was, however, found to have construct validity and showed significant improvement with subsequent repetitions.

There are a few other identifiable limitations of training on the cadaver model compared to in vivo training. The most important one is the lack of bleeding on the cadaver model. Hence, the scenario of “iatrogenic bleeding control” cannot be created to make the trainees deal with bleeding disasters. Work is underway to overcome this limitation.

The cost and legal constraints make cadaveric training a very precious resource. The challenge for future expansion of FFC training rests on the limited number of licensed cadaver training centers. The cost of cadavers for this study was reduced by ensuring their use for a range of disciplines, including colorectal, orthopedics, otolaryngology, cardiovascular, and urology training courses. The creation of a simulated appendectomy model using the small bowel, as shown in our study, ensures that the task is repeatable and integrates the practice of several basic skills within a full task. Such initiatives can ensure maximum use of cadavers.

## CONCLUSION

Our study demonstrates that the use of fresh human cadavers is a very effective model to train junior trainees in basic laparoscopic skills up to a safe level within a short time period. Valid assessment scales can successfully differentiate between novice and expert surgeons on the fresh human cadaver model, suggesting construct validity.

## References:

1. Holland J, Waugh L, Horgan A, Paleri V, Deehan D. Cadaveric hands-on training for surgical specialties: Is this back to the future for surgical skills development? *Jr Surg Educ.* 2010;68:110–116.
2. Verdaasdonk EG, Stassen LP, Monteny IJ, Dankelman J. Validation of a new basic virtual reality simulator for training of basic endoscopic skills. *Surg Endosc.* 2006;20:511–518.
3. Gallagher AG, McClure N, McGuigan J, Crothers I, Browning J. Virtual reality training in laparoscopic surgery: a preliminary assessment of Minimally Invasive Surgical Trainer Virtual Reality (MIST VR). *Endoscopy.* 1999;31:310–313.
4. Gallagher AG, Lederman AB, McGlade K, Satava RM, Smith CD. Discriminative validity of the Minimally Invasive Surgical

Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance. *Surg Endosc.* 2004;18:660–665.

5. Schijven M, Jakimowicz J. Construct validity: experts and novices performing on Xitact LS500 laparoscopy simulator. *Surg Endosc.* 2003;17:803–810.

6. Anastakis DJ, Regehr G, Reznick RK, et al. Assessment of technical skills transfer from the bench training model to the human model. *Am J Surg.* 1999;177:167–170.

7. Levine RL, Kives S, Cathey G, et al. The use of lightly embalmed (fresh tissue) cadavers for resident laparoscopic training. *J Minim Invasive Gynecol.* 2006;13:451–456.

8. Gordinier ME, Granai CO, Jackson ND, Metheny WP. The effects of a course in cadaver dissection on resident knowledge of pelvic anatomy: an experimental study. *Obstet Gynecol.* 1995;86:137–139.

9. Cundiff GW, Weidner AC, Visco AG. Effectiveness of laparoscopic cadaveric dissection in enhancing resident comprehension of pelvic anatomy. *J Am Coll Surg.* 2001;192:492–497.

10. Ritter EM, Scott DJ. Design of a Proficiency Based Skills Training Curriculum for the Fundamentals of Laparoscopic Surgery. *Surg Innov.* 2007;14:107–112.

11. Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomised controlled trial. *Surg Endosc.* 2002;16:957–961.

12. Moulton CA, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect?: a randomised controlled trial. *Ann Surg.* 2006;244:400–409.

13. Gumbs AA, Hogle NJ, Fowler DL. Evaluation of resident Laparoscopic performance using global assessment of laparoscopic skills. *J Am Coll Surg.* 2006;204:308–313.

14. 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.

15. Vassiliou MC, Feldman LS, Fraser SA, et al. Evaluating intraoperative laparoscopic skills; Direct observation versus Blinded videotaped performances. *Surg Innov.* 2007;14:211–216.

16. Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A. Development of a virtual reality training curriculum for laparoscopic cholecystectomy. *Br J Surg.* 2009;96:1086–1093.

17. Babineau TJ, Becker J, Gibbons G, et al. The cost of operative training for residents. *Arch Surg.* 2004;139:366–370.

18. Bridges M, Diamond DL. The financial impact of teaching surgical residents in the operating room. *Am J Surg.* 1999;177:28–32.

19. Levine RL, Kives S, Cathey G, et al. The use of lightly embalmed (fresh tissue) cadavers for resident laparoscopic training. *J Minim Invasive Gynecol.* 2006;13:451–456.

20. Eversbusch A, Grantcharov TP. Learning curves and impact of psychomotor training on performance in simulated colonoscopy; a randomised trial using a virtual reality endoscopy trainer. *Surg Endosc.* 2004;18:1514–1518.

21. Sharma M, Hance J, Horgan A. Comparison of fresh frozen cadavers and high fidelity virtual reality simulator as methods of laparoscopic training. *Br J Surg.* 2011;98:Suppl I.

22. Aggarwal R, Grantcharov TP, Eriksen JR, et al. An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg.* 2006;244:310–314.