An evaluation of the endoscopic surgical skills assessment using a video analysis software program

Takahisa Suzuki · Hiroyuki Egi · Minoru Hattori ·
Masakazu Tokunaga · Hiroyuki Sawada ·
Hideki Ohdan

Abstract

Backgrounds   Recently, the endoscopic surgical skills of surgeons have become a focus of many investigators and society as a whole. Hence, the importance of the training and education programs used to ensure safer endoscopic surgery has been gradually increasing. We developed our own motion analysis system called the HUESAD, and already demonstrated its construct validity. In this study, we verified that another motion analysis system (which uses the Dartfish Software program) could assess surgeons' endoscopic surgical skills.

Methods   Experts (who had performed more than 100 laparoscopic surgeries) and novices (who had no experience performing laparoscopic surgery) were recruited for this study. The task was suturing in a dry box trainer. The time and the locus tracing of both sides' needle holders were analyzed using the new video analysis system (Dartfish Software).

Results   There were statistically significant differences between the experts and novices in all three variables assessed (task time: $p = 0.0011$, the locus tracing of the left sides' needle holders: $p = 0.0011$, the locus tracing of the right sides' needle holders: $p = 0.0011$).

Conclusions   The results of this study demonstrated that the results of the motion analysis by the Dartfish Software were well correlated with the surgeon's skill level.

Keywords   Endoscopic surgery · Objective skill assessment · HUESAD · Suturing movement · Motion analysis · Dartfish Software

Laparoscopic surgery has spread quickly since the first report of a laparoscopic cholecystectomy in 1987 [1], and is now one of the standard approaches to various types of surgery. Clinically, the curability of laparoscopic surgery must be same as in open surgery, and safety and scar cosmetics are also important. The establishment of training and education systems for laparoscopic surgery is necessary, because it will help surgeons master the new skills in a safe and effective way [2]. The Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and the European Association of Endoscopic Surgery (EAES) established minimum requirements for those performing laparoscopic surgery [3, 4]. Research concerning the assessment of endoscopic surgical skills has been developing rapidly because so many people are interested in performing endoscopic surgery.

The objective structured assessment of technical skills (OSATS) consists of six stations where participants perform tasks on live animal or bench models. Their performance is assessed using checklists and a global rating scale [5]. This method was the first reliable system for assessing surgical skills [6], but this method requires many people and a long time to perform. Therefore, some new assessment tools have been developed, and their construct validity has been demonstrated.

Surgical simulation is an effective tool for training and assessment. One of the key points in a training program is

T. Suzuki · H. Egi · M. Tokunaga · H. Sawada · H. Ohdan
Department of Gastroenterological and Transplant Surgery,
Applied Life Sciences, Institute of Biomedical & Health
Sciences, Hiroshima University, 1-2-3 Kasumi, Minami-ku,
Hiroshima 734-8551, Japan

H. Egi (✉) · M. Hattori
Advanced Medical Skills Training Center, Institute of
Biomedical & Health Sciences, Hiroshima University,
Hiroshima, Japan
e-mail: hiroegi@yahoo.co.jp
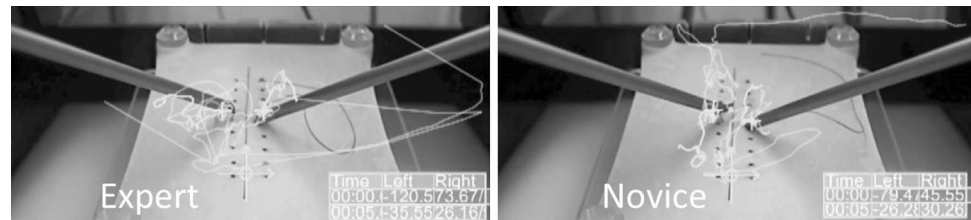
providing trainees with immediate feedback of their performance. Simulators can produce learning curves outside the operating theater in a pressure-free environment, without requiring formal supervision. Virtual reality simulators (MIST-VR [7, 8], Lap Mentor [9, 10], Lap Sim [11, 12], ProMIS [13], etc.) and motion analysis systems (IC-SAD [14–16], Adept [17, 18], HUESAD, etc.) are now able to provide real-time scores in the form of evidence-based reports (path length, errors, overall score, time, etc.). These methods usually require two phases: (a) training and (b) classification. During the training phase, representative data from each surgical level (e.g., novices, intermediates, experts) is used to establish the different class representative of the competence level. In the classification phase, the data recovered for a new surgeon are compared to those classes, and an assignation to one of them is performed based on a likelihood probability.

We developed our own motion analysis system for endoscopic surgery, called the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD) [19]. The HUESAD has reliable parameters, visual-spatial ability, smoothness and accuracy, and we showed construct validity of the HUESAD [19–22]. Moreover, we found that there was strong correlation between the motion analysis in the HUESAD assessment and the OSATS checklist and the global rating score [23]. The HUESAD has been demonstrated to be reliable as a system for the objective assessment of endoscopic surgical skills.

However, the HUESAD is a motion analysis system in dry box trainer, which means that it is used outside the operating theater.

We therefore used a new video analysis software program (Dartfish Software) that enables the user to take measurements such as the angles, distances and timing directly on digital video recordings made during a procedure, and offers the potential to decrease the bias associated with subjective image assessments. The Dartfish Software program has previously been used to develop performance-enhancing sports video training applications and exclusive televised broadcast footage.

The aim of the present study was to verify whether the video analysis software program (Dartfish Software) could be used to assess surgeons' endoscopic surgical skills.

Methods

Box trainer

A box trainer (Endowork Pro II; MC Medical, Tokyo, Japan) was used in this study. A television image showing the inside of the box was visualized through a CCD camera on a 16-in monitor (Sharp Electronics, Osaka, Japan). Laparoscopic instruments (KOH Macro Needle Holder, 45 cm; Karl Storz, Tuttlingen, Germany) and surgical thread (3-0 braided silk SH-1; Ethicon, Somerville, NJ) were used.

Suturing tasks using a box trainer

The suturing task was evaluated in this study. This task is suturing five times without knot tying in a box trainer. The time and the locus tracing of the needle holders on both sides were analyzed using the video analysis software program (Dartfish Software). This software program allowed the time and the locus tracing to be measured for each subject (see Fig. 1).

Study design

Six experts (who had performed more than 100 laparoscopic surgeries) and eleven novices (who had no experience performing laparoscopic surgery) were recruited for this study. The aim of the task was to simply add five sutures in dry box trainer. The time and the locus tracing of the needle holders on both sides were analyzed using the Dartfish Advanced Video Analysis Software Program (version 5.5, Dartfish, Fribourg, Switzerland). This video measurement software de-interlaces video files to 30 images per second (sampling rate ~ 0.033 s). A previous study demonstrated that using the Dartfish Advanced Video Analysis Software program is an efficient approach to improve the reliability of visual video assessments (the use of video analysis software increased the interpreter reliability of video gait assessments in children with cerebral palsy).

Statistical analyses

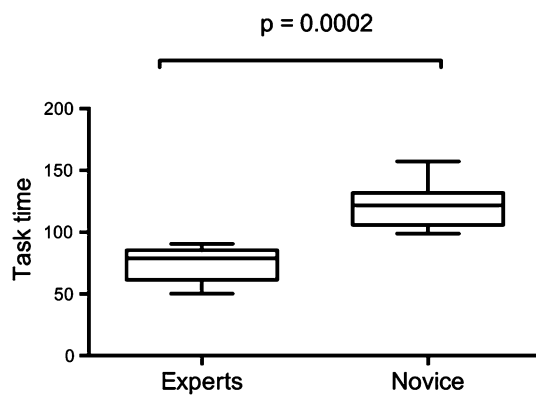All data were processed and analyzed using the SPSS software package (ver. 19.0, SPSS Inc., Chicago, IL). The

Fig. 2 Total execution time to perform the task

Table 1 Our classification methods can correctly classify 80 % of experts in the two groups (cross-validation test)

Actual	Predicted	
	Expert	Novice
Expert	6	0
Novice	1	10

Wilcoxon signed-ranks test for related data was used to assess the differences in performance for all of the parameters measured for each group. The Mann–Whitney U-test was applied for comparisons among each of the training groups. The study used a linear discriminant analysis (LDA), a classification method that can automatically determine the threshold level for classifying experts or novices according to the time spent and the locus tracing of the needle holders on both sides. The performance of the classification methods was examined using a cross-validation. Statistical significance was defined as a p value <0.05.

Results

The results indicated that there was a statistically significant difference between the experts and novices in all three variables examined (task time: $p = 0.0011$, the locus tracing of the left sides' needle holders: $p = 0.0011$, the locus tracing of the right sides' needle holders: $p = 0.0011$) (Figs. 2, 3).
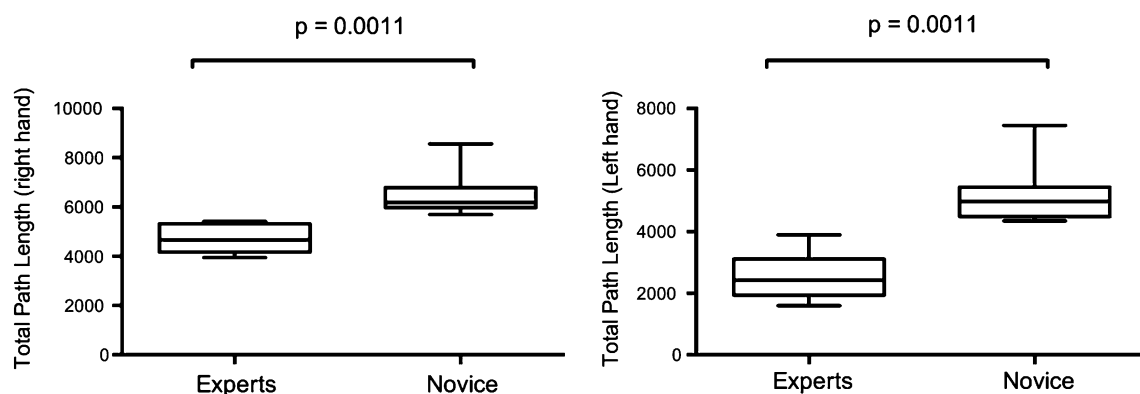
The best LDA results were obtained for the combination of the three parameters together (the time and the locus tracing of the needle holders on both sides). Table 1 shows the results of this classification in a confusion matrix. In this matrix, the actual (ground truth) classifications are in the rows, and the classifications predicted by the LDA are in the columns. Our classification methods could therefore correctly classify 100 % of the experts and novices.

Discussion

The results of this study demonstrate that the time and the locus tracing to complete a task were well correlated with the operator's skill level. The "time to complete a task" has traditionally been used to assess technical performance [24]. However, faster performance is not always associated with better quality and improved outcomes. Locus tracing may better reflect the motor accuracy.

New tools have emerged over the last few years for the objective assessment of technical performance [20, 25]. A motion analysis allows for an assessment of the surgical dexterity using parameters that are extracted from the movement of the hands or laparoscopic instruments. An objective assessment of laparoscopic skill can be carried out using a motion analysis if endpoints for each parameter are quantified according to pre-defined levels of experience. The conversion of motion analysis data into competency-based scores or indices could provide a valuable source for trainee feedback. Such feedback can be useful,

Fig. 3 Total path length (right hand and left hand) between the expert and novice group

because it provides a quantitative index to define varying levels of experience, which trainees can work toward [26].

The HUESAD, our previously developed system which can precisely trace the movement of the tip of a laparoscopic instrument, was developed for the objective assessment of technical performance. Previous studies demonstrated the construct validity and reliability of the HUESAD [19, 21–23, 27]. The HUESAD can define three different parameters: the integrated deviation (the visual-spatial ability) [19], the peak velocity (smoothness) [21] and the approach time (accuracy) [27]. By these three parameters, the HUESAD can provide feedback not only on the skill level of the trainee, but also provide data about the weak points. The HUESAD can be used to assess basic training as the subjects utilize simple orientation and movement skills in a non-anatomical environment. We have evaluated the correlation between the HUESAD and OSATS scores in a concurrent validity study [23]. However, it remains unknown whether the HUESAD can distinguish between different levels of performance in the operating theater.

The purpose of this study was to verify whether a video analysis software program (the Dartfish Software) could assess surgeons' endoscopic surgical skills in a real operation. The Dartfish Software is not only used in sports science [6, 25, 28], but also in situations associated with medical support, such as providing instant visual feedback in treatment sessions, which allows patients to immediately see exactly what their body is doing.

On the other hand, the Dartfish software program can also assess the total endoscopic surgical skills using a simple method, which may be convenient to evaluate a surgeon's skills in the clinical setting. In terms of training, the Dartfish Software program can assess the degree of skills attainment, and the HUESAD can assess the surgeon's skills in detail, and can provide information such as "what are the surgeon's weaknesses?"

In this study, we assessed a simple task: suturing in dry box trainer. This study showed the system's construct validity, and the Dartfish Software could differentiate between levels of performance, but it cannot be used for formative assessment. In this study, locus tracing was calibrated using conventional two-dimensional (2D) data, so it was not the actual measured value, indicating that this data had limited accuracy.

Conclusion

The results of this study demonstrated that the results of the motion analysis using the Dartfish Software program were well correlated with a surgeon's skill level. The construct validity was supported by the results of this study. In addition, the Dartfish Software program is simple to use, but has lower accuracy than real locus tracing. Further development of the system will be necessary to provide more relevant data, including work in three-dimensions (3D) and in the operating theater.

References

1. Litynski GS (1999) Profiles in laparoscopy: Mouret, Dubois, and Perissat: the laparoscopic breakthrough in Europe (1987–1988). JSLS 3(2):163–167
2. Botden SM, Buzink SN, Schijven MP, Jakimowicz JJ (2007) Augmented versus virtual reality laparoscopic simulation: what is the difference? A comparison of the ProMIS augmented reality laparoscopic simulator versus LapSim virtual reality laparoscopic simulator. World J Surg 31(4):764–772. doi:10.1007/s00268-006-0724-y
3. Guidelines for granting of privileges for laparoscopic and/or thoracoscopic general surgery. Society of American Gastrointestinal Endoscopic Surgeons (SAGES) (1998). Surg Endosc 12(4):379–380
4. Training and assessment of competence (1994). Surg Endosc 8(6):721–722
5. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 84(2):273–278
6. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, Hutchison C (1999) Assessment of technical skills transfer from the bench training model to the human model. Am J Surg 177(2):167–170
7. Gallagher AG, Lederman AB, McGlade K, Satava RM, Smith CD (2004) Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance. Surg Endosc 18(4):660–665. doi:10.1007/s00464-003-8176-z
8. Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D (2006) Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? Surg Endosc 20(1):104–112. doi:10.1007/s00464-005-0054-4
9. Zhang A, Hunerbein M, Dai Y, Schlag PM, Beller S (2008) Construct validity testing of a laparoscopic surgery simulator (Lap Mentor): evaluation of surgical skill with a virtual laparoscopic training simulator. Surg Endosc 22(6):1440–1444. doi:10.1007/s00464-007-9625-x
10. Yamaguchi S, Konishi K, Yasunaga T, Yoshida D, Kinjo N, Kobayashi K, Ieiri S, Okazaki K, Nakashima H, Tanoue K, Maehara Y, Hashizume M (2007) Construct validity for eye-hand coordination skill on a virtual reality laparoscopic surgical simulator. Surg Endosc 21(12):2253–2257. doi:10.1007/s00464-007-9362-1
11. Duffy AJ, Hogle NJ, McCarthy H, Lew JI, Egan A, Christos P, Fowler DL (2005) Construct validity for the LAPSIM laparoscopic surgical simulator. Surg Endosc 19(3):401–405. doi:10.1007/s00464-004-8202-9
12. Woodrum DT, Andreatta PB, Yellamanchilli RK, Feryus L, Gauger PG, Minter RM (2006) Construct validity of the LapSim laparoscopic surgical simulator. Am J Surg 191(1):28–32. doi:10.1016/j.amjsurg.2005.10.018

13. Pellen MG, Horgan LF, Barton JR, Attwood SE (2009) Construct validity of the ProMIS laparoscopic simulator. Surg Endosc 23(1):130–139. doi:10.1007/s00464-008-0066-y

14. Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A (2002) Relationship between skill and outcome in the laboratory-based model. Surgery 131(3):318–323

15. Bann SD, Khan MS, Darzi AW (2003) Measurement of surgical dexterity using motion analysis of simple bench tasks. World J Surg 27(4):390–394. doi:10.1007/s00268-002-6769-7

16. Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A (2004) Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. Surg Endosc 18(11):1608–1612. doi:10.1007/s00464-003-9312-5

17. Francis NK, Hanna GB, Cuschieri A (2002) The performance of master surgeons on the Advanced Dundee Endoscopic Psychomotor Tester: contrast validity study. Arch Surg 137(7):841–844

18. Macmillan AI, Cuschieri A (1999) Assessment of innate ability and skills for endoscopic manipulations by the Advanced Dundee Endoscopic Psychomotor Tester: predictive and concurrent validity. Am J Surg 177(3):274–277

19. Egi H, Okajima M, Yoshimitsu M, Ikeda S, Miyata Y, Masugami H, Kawahara T, Kurita Y, Kaneko M, Asahara T (2008) Objective assessment of endoscopic surgical skills by analyzing direction-dependent dexterity using the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD). Surg Today 38(8):705–710. doi:10.1007/s00595-007-3696-0

20. Egi H, Tokunaga M, Hattori M, Ohdan H (2012) Evaluating the correlation between the HUESAD and OSATS scores: concurrent validity study. Minim Invasive Ther Allied Technol. doi:10.3109/13645706.2012.742113

21. Egi H, Okajima M, Kawahara T, Yoshimitsu M, Sumitani D, Tokunaga M, Takeda H, Itamoto T, Ohdan H (2010) Scientific assessment of endoscopic surgical skills. Minim Invasive Ther Allied Technol 19(1):30–34. doi:10.3109/13645700903493017

22. Tokunaga M, Egi H, Hattori M, Yoshimitsu M, Sumitani D, Kawahara T, Okajima M, Ohdan H (2012) Approaching time is important for assessment of endoscopic surgical skills. Minim Invasive Ther Allied Technol. doi:10.3109/13645706.2011.596547

23. Rucci JA, Tomporowski PD (2010) Three types of kinematic feedback and the execution of the hang power clean. J Strength Cond Res 24(3):771–778. doi:10.1519/JSC.0b013e3181cbab96

24. Xafis V, Babidge W, Field J, Altree M, Marlow N, Maddern G (2013) The efficacy of laparoscopic skills training in a mobile simulation unit compared with a fixed site: a comparative study. Surg Endosc. doi:10.1007/s00464-013-2798-6

25. Padulo J, Annino G, Migliaccio GM, D'Ottavio S, Tihanyi J (2012) Kinematics of running at different slopes and speeds. J Strength Cond Res 26(5):1331–1339. doi:10.1519/JSC.0b013e318231aafa

26. Mason JD, Ansell J, Warren N, Torkington J (2013) Is motion analysis a valid tool for assessing laparoscopic skill? Surg Endosc 27(5):1468–1477. doi:10.1007/s00464-012-2631-7

27. Tokunaga M, Okajima M, Egi H, Yoshimitsu M, Sumitani D, Kawahara T, Hattori M, Ohdan H (2010) The importance of stressing the use of laparoscopic instruments in the initial training for laparoscopic surgery using box trainers: a Randomized Control Study. J Surg Res. doi:10.1016/j.jss.2010.11.906

28. Panait L, Akkary E, Bell RL, Roberts KE, Dudrick SJ, Duffy AJ (2009) The role of haptic feedback in laparoscopic simulation training. J Surg Res 156(2):312–316. doi:10.1016/j.jss.2009.04.018