

Objective evaluation of expert performance during human robotic surgical procedures

Timothy N. Judkins · Dmitry Oleynikov ·
Nick Stergiou

Received: 23 October 2007 / Accepted: 5 December 2007 / Published online: 4 January 2008
© Springer London 2007

Abstract Robotic laparoscopic surgery has revolutionized minimally invasive surgery and has increased in popularity due to its important benefits. However, evaluation of surgical performance during human robotic laparoscopic procedures in the operating room is very limited. We previously developed quantitative measures to assess robotic surgical proficiency. In the current study, we want to determine if training task performance is equivalent to performance during human surgical procedures performed with robotic surgery. An expert with more than 5 years of robotic laparoscopic surgical experience performed two training tasks (needle passing and suture tying) and one human laparoscopic procedure (Nissan fundoplication) using the da VinciTM Surgical System (dVSS). Segments of the human procedure that required needle passing and suture tying were extracted. Time to task completion, distance traveled, speed, curvature, and grip force were measured at the surgical instrument tips. Single-subject analysis was used to compare training task

performance and human surgical performance. Nearly all objective measures (8 out of 13) were significantly different between training task performance and human surgical performance for both the needle passing and the suture tying tasks. The surgeon moved slower, made more curved movements, and used more grip force during human surgery. Even though it appears that the surgeon performed better in the training tasks, it is likely that during human surgical procedures, the surgeon is more cautious and meticulous in the movements performed in order to prevent tissue damage or other complications. The needle passing and the suture tying training tasks may be suitable to establish a foundation of surgical skill; however, further training may be necessary to improve transfer of learning to the operating room. We recommend that more realistic training tasks be developed to better predict performance during robotic surgical procedures and testing the transferability of basic skill acquisition to surgical performance.

Keywords Nissan fundoplication · Operating room · Performance · Quantitative · Robotic surgery

T. N. Judkins (✉)
U.S. Army Research Laboratory, ATTN: AMSRD-ARLHR-SC,
Aberdeen Proving Ground, Aberdeen, MD 21005, USA
e-mail: tjudkins@gmail.com

D. Oleynikov
Department of Surgery, University of Nebraska Medical Center,
Omaha, NE, USA

N. Stergiou
HPER Biomechanics Laboratory,
University of Nebraska at Omaha, Omaha, NE, USA

N. Stergiou
Department of Environmental,
Agricultural and Occupational Health Sciences,
College of Public Health, University of Nebraska Medical
Center, Omaha, NE, USA

Introduction

Previous studies have evaluated manual and robotic laparoscopy and examined surgical performance during training tasks and during animal procedures [1–13]. Training tasks in these studies included needle passing and suture tying. Animal procedures included anastomosis and cholecystectomy. However, evaluation of surgical performance during human robotic laparoscopic procedures in the operating room has not been investigated previously.

Regarding robotic laparoscopy, we previously identified a number of objective variables, including speed, path

curvature, and grip force, that can be used to quantify surgical performance [7, 14, 15]. These previous studies were limited in terms of their external validity because actual human robotic surgical procedures were not evaluated. Thus, their results were inferred to the actual procedures. When human robotic surgical procedures have been observed, performance measures were limited to time of completion or surgeon opinion [16–18]. Thus, there has been no quantification of performance during human robotic surgical procedures in the operating room. Therefore, it is necessary to observe robotic surgical procedures in the operating room to determine if objective measures or performance during human procedures are equivalent to training task performance.

Therefore, the purpose of this study was to determine if training task performance is equivalent to performance during human surgical procedures performed with robotic surgery. We monitored performance of an expert in robotic laparoscopic surgery during two training tasks and one human procedure (Nissan fundoplication). Movement kinematics was recorded via the da VinciTM Surgical System (dVSS) and custom data logging software. Objective measures of performance, that have been previously validated to discern novice and expert performance, were computed. We hypothesized that performance between training and human surgical procedures is similar. Specifically, we hypothesized that objective measures of performance will not be significantly different between training tasks and human procedures.

Methods and materials

Experimental protocol

An expert with more than 5 years of robotic laparoscopic surgical experience performed two training tasks and one human laparoscopic procedure using the da VinciTM Surgical System (dVSS). The expert performed five trials of two training tasks (needle passing and suture tying) during two sessions for a total of ten trials of each task, while kinematics of the surgical instrument tips was recorded from the dVSS. In addition, the expert performed a Nissan fundoplication. This procedure was selected because it requires needle passing and suture tying. The surgical procedure was observed in the operating room, while kinematics of the surgical instrument tips was recorded from the dVSS via an Ethernet connection. Since data were recorded passively, analyzed offline, and no patient identifiers were used, this study was considered exempt by the Institutional Review Board of University of Nebraska Medical Center. Verbal consent was obtained from the surgeon.

Training tasks

The expert performed two tasks using the dVSS: needle passing (NP) and suture tying (ST). The needle passing task required passing a 26 mm surgical needle through six holes in a latex tube (see Fig. 1a). The expert started from the proximal holes and proceeded in order to the distal. The suture tying task required tying two knots with a 100 mm × 0.5 mm surgical suture using the intracorporeal knot (see Fig. 1b). The expert performed the tasks by manipulating the dVSS from the surgeon's console (see Fig. 1c).

Surgical procedures

One surgical procedure (Nissan fundoplication) was observed for this study. A Nissan fundoplication is wrapping of the stomach around the esophagus. This procedure is a common gastrointestinal procedure that is performed with robotic surgery. This surgical procedure requires similar movements to the training tasks (needle passing and suture tying).

Data analysis

Objective performance measures were calculated for each trial for the training tasks and for segments of the human procedure. Segments of the human procedure were identified by the portions that required needle passing and suture tying. Four needle passing segments and four suture tying segments were identified during the human procedure. Additionally, since each needle passing segment during the human procedure required passing the needle through tissue twice, objective measures for the needle passing task were calculated for the first two passes during each training task trial.

All objective measures of performance were based on kinematics and kinetics of the instrument tips of the dVSS. Kinematics of the dVSS was collected using the Application Programmer's Interface (API) provided by Intuitive Surgical, Inc. (Sunnyvale, CA). A custom LabView (National Instruments, Inc., Concord, MA) program was written to interface to the dVSS via an Ethernet connection. Data was streamed at approximately 75 Hz (determined by the API). All post-processing of data was performed in MATLAB 6.5 (Mathworks, Inc., Natick, MA). All kinematic data was down-sampled to 5 Hz using a cubic spline to enforce a constant sampling rate between data points. Variables of interest, streamed from the API, were position (x, y, and z location) of the right and left instrument tips and the grip force applied by the left and right grasper.

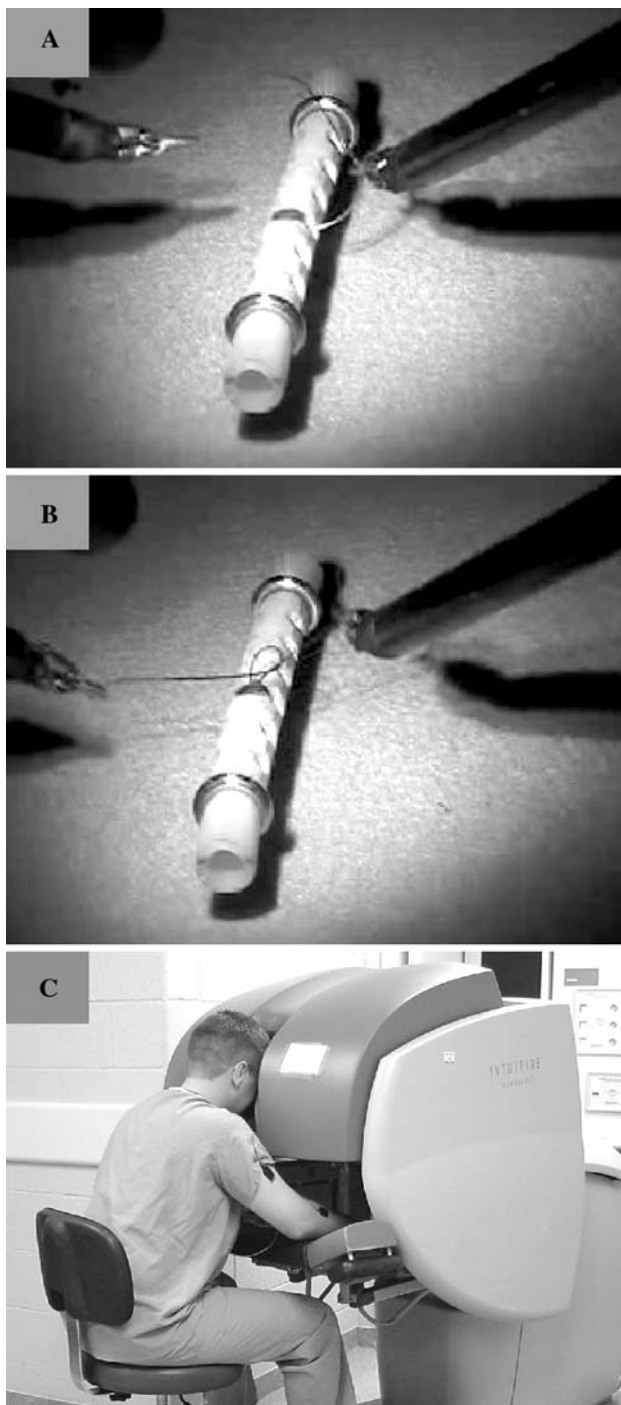


Fig. 1 Experiment Setup. **a** Needle Passing. **b** Suture Tying. **c** Subject seated at surgeon's console of dVSS

The measurements computed from the robot kinematics were position, velocity, and acceleration of the instrument tips. All measurements were calculated for both the left and the right instrument (or grasper). Velocity and acceleration of the instrument tips were directly calculated by computing the first and second derivatives, respectively, of the

positions of the instrument tips. The kinematic measures (also referred here as objective performance variables) used, were time to task completion (TTC), total distance traveled (D), speed (S), curvature (κ), and grip force (F). Time to task completion is the time required to complete a given training task or human surgery segment. For the training tasks, start and end time were identified as the time when the instrument tips were within 1 cm of the starting positions. For the human procedure segments, start and end time were identified manually using a video recording that was synchronized to the kinematic and kinetic data. Total distance traveled is the sum of Euclidean distances between each time sample. Speed is calculated as the magnitude of the velocity. The mean (S_{mean}) and standard deviation (S_{std}) of speed were computed for each training task trial and each segment of the human procedure.

Curvature measures the straightness of the path and is calculated at each point on the path by the following equation [19, 20]:

$$\kappa = \frac{|\dot{r} \times \ddot{r}|}{|\dot{r}|^3} \quad (1)$$

where \dot{r} is the velocity of a point r on the three-dimensional path and \ddot{r} is the acceleration of point r . The median (κ_{med}) and 95% confidence interval (κ_{CI}) were computed for each training task trial and each segment of the human procedure. The 95% confidence interval was computed as defined by Campbell and Gardner [21].

Grip force (F) was provided by the dVSS API and represented a percentage of the maximum torque of the servos that drive the graspers. In order to verify the linearity of the grip force, a force sensing resistor (FSR) was squeezed while measurements from the dVSS and FSR were collected simultaneously. The resistance of the FSR is proportional to the force applied; therefore, grip force could be directly measured. dVSS and FSR grip force measurements were compared using a linear regression fit. Right and left grip force measured by the dVSS were strongly correlated ($R^2 = 0.97$ and 0.91 , respectively) with FSR measurements. Mean grip force (F_{mean}) was computed for each training task trial and each segment of the human procedure.

Statistical analysis

Since only one expert participated in this study, a single subject analysis technique (Model Statistic; [22–24]) was used to compare the means of training task performance (TRAIN) to human procedure (HUMAN) for each objective measure and each task at $\alpha = 0.05$ level. In this single subject analysis, the difference between the means of each condition is compared with the product of the mean

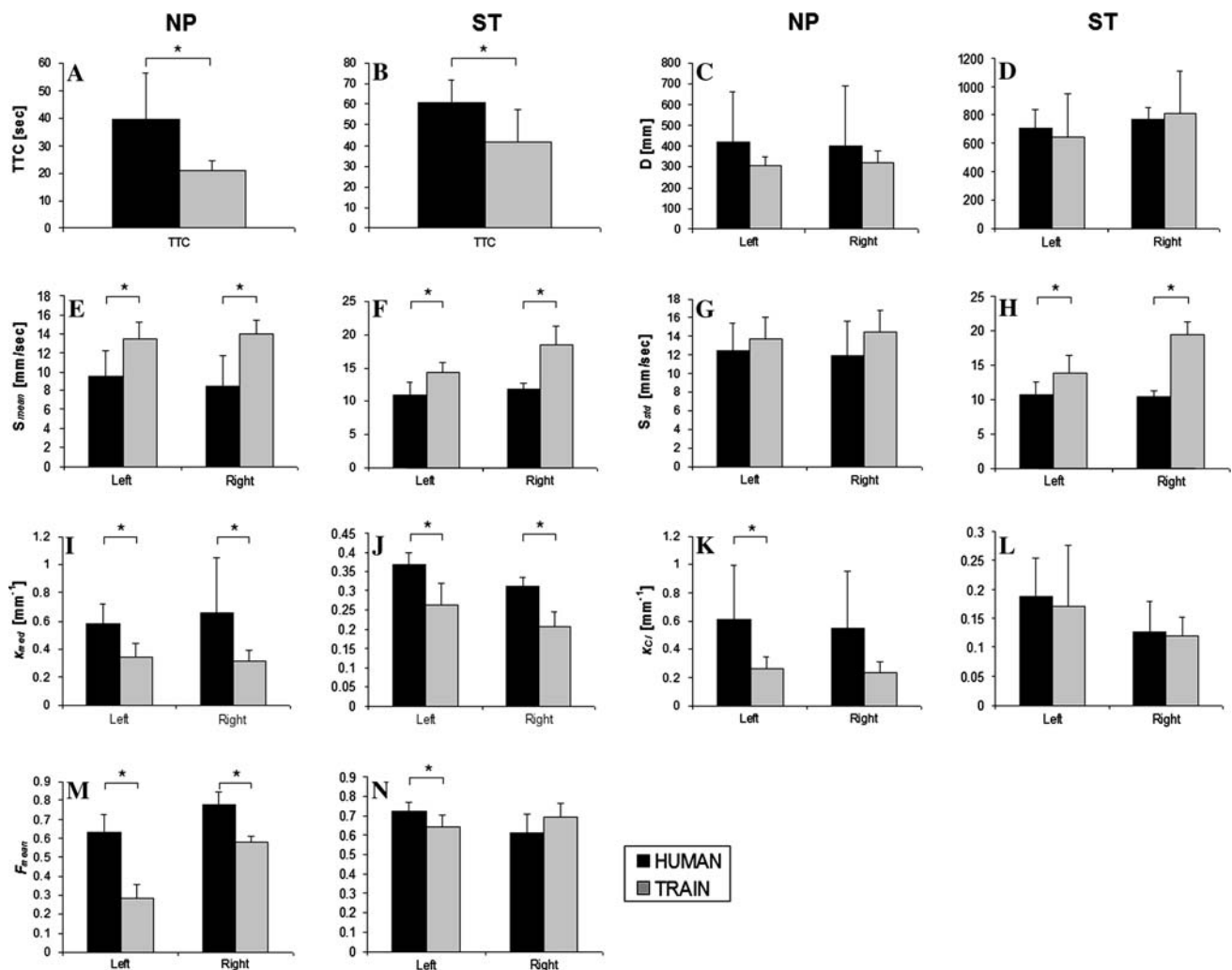


Fig. 2 Objective measures comparing training performance (*TRAIN*) to human surgical performance (*HUMAN*) during needle passing (*NP*) and suture tying (*ST*) movements for TTC (time to task completion; **a**, **b**), right and left *D* (distance traveled; **c**, **d**), right and left S_{mean} (mean speed; **e**, **f**), right and left S_{std} (standard deviation of speed; **g**, **h**), right

standard deviation and a criterion test statistic based on the number of trials [24].

Results

Needle passing

Except *D*, S_{std} , and right κ_{CI} , all other objective measures (8 out of 13 comparisons; 61.5%) were significantly different between *TRAIN* and *HUMAN* performance at $\alpha = 0.05$ level (Fig. 2). TTC was significantly larger during *HUMAN* performance as compared to *TRAIN* performance. Right and left S_{mean} were significantly smaller during *HUMAN* performance as compared to *TRAIN* performance. Right and left κ_{med} were significantly larger during *HUMAN* performance as compared to *TRAIN* performance. Left κ_{CI}

and left κ_{med} (median curvature; **i**, **j**), right and left κ_{CI} (confidence interval of curvature; **k**, **l**), and right and left F_{mean} (mean grip force; **m**, **n**). Lines with stars above them indicate a significant difference at $\alpha = 0.05$ level between *TRAIN* and *HUMAN* for the variable indicated

was significantly larger during *HUMAN* performance as compared to *TRAIN* performance. Right and left F_{mean} were significantly larger during *HUMAN* performance as compared to *TRAIN* performance.

Suture tying

Except *D*, κ_{CI} , and right F_{mean} , all objective measures (8 out of 13 comparisons; 61.5%) were significantly different between *TRAIN* and *HUMAN* performance at $\alpha = 0.05$ level (Fig. 2). TTC was significantly larger during *HUMAN* performance as compared to *TRAIN* performance. Right and left S_{mean} were significantly smaller during *HUMAN* performance as compared to *TRAIN* performance. Right and left S_{std} were significantly smaller during *HUMAN* performance as compared to *TRAIN*

performance. Right and left κ_{med} were significantly larger during HUMAN performance as compared to TRAIN performance. Left F_{mean} was significantly larger during HUMAN performance as compared to TRAIN performance.

Discussion

Our results refuted our hypothesis, since we found that robotic surgical performance during human surgery is not similar to performance during training for the two tasks selected, suture tying and needle passing. The surgeon moved slower, made more curved movements, and used more grip force during the human surgery. Even though it appears that the surgeon performed better in the training tasks, it is likely that during the human surgical procedure, the surgeon is more cautious and meticulous in the movements performed in order to prevent tissue damage or other complications. Furthermore, space limitations inside the patient's body during the human surgery can also affect performance.

Furthermore, it is possible that our results are task dependent and both the needle passing and the suture tying training tasks cannot accurately represent performance during the human surgery. Therefore, additional training tasks are needed to be investigated to identify if they can better simulate surgical procedures in order to predict human surgical performance. Better surgical simulation can be accomplished using simulated tissue for training tasks, performing training tasks inside a simulated body cavity, or by using virtual reality. Virtual reality has proved to be effective for manual laparoscopy for improving performance in the operating room. In a randomized, double-blinded study, researchers found that virtual reality training for manual laparoscopy significantly improves performance during human gall bladder dissection [25]. Further investigation is needed to determine if virtual reality is an effective training paradigm for robotic surgery with strong transfer of learning to the operating room.

Training tasks are important in that they provide a foundation of basic skills for human surgical procedures [1–11, 14, 15]. In manual laparoscopy, training with basic skills has been shown to improve performance during animal procedures. The needle passing and the suture tying training tasks may be suitable to establish a foundation of surgical skill; however, further training with more advanced tasks may be necessary to improve transfer of learning to the operating room. Future work is needed to determine if novices can transfer basic surgical skills learned during training tasks to surgical procedures in animal models before objective skill assessment during

training task can be used as a measure of surgical proficiency.

In conclusion, our study showed that performance during simple training tasks such as suture tying and needle passing is not equivalent to performance during human robotic surgical procedures. The surgeon in this study was more cautious and meticulous during human procedures compared to training as indicated by reduced speed and more curved movements. While training tasks are important for basic skill acquisition, training task performance does not correlate with human procedure performance. We recommend that more realistic and advanced training tasks be developed to better predict performance during human robotic surgical procedures and to test the transferability of skill acquisition to surgical performance.

Acknowledgments This work was supported by NIH (K25HD047194), NIDRR (H133G040118), and the Nebraska Research Initiative.

Appendix: Abbreviations and acronyms

General

API	Application Programmer's Interface
dVSS	da Vinci TM Surgical System
FSR	Force sensing resistor

Tasks

NP	Needle passing
ST	Suture tying

Conditions

TRAIN	Training tasks
HUMAN	Human laparoscopic procedure in operating room

Objective measures of performance

TTC	Time to task completion
D	Distance traveled
S_{mean}	Mean speed
S_{std}	Standard deviation of speed
κ_{med}	Median path curvature
κ_{CI}	Path curvature confidence interval
F_{mean}	Mean grip force

References

1. Yohannes P, Rotariu P, Pinto P, Smith AD, Lee BR (2002) Comparison of robotic versus laparoscopic skill: is there a difference in the learning curve? *Urology* 60:39–45

2. Sarle R, Tewari A, Shrivastava A, Peabody J, Menon M (2004) Surgical robotics and laparoscopic training drills. *J Endourol* 18(1):63–67
3. Prasad SM, Maniar HS, Soper NJ, Damiano RJ, Klingensmith ME (2002) The effect of robotic assistance on learning curves for basic laparoscopic skills. *Am J Surg* 183(6):702–707
4. Moorthy K, Munz Y, Dosis A, Hernandez J, Martin S, Bello F, Rockall T, Darzi A (2004) Dexterity enhancement with robotic surgery. *Surg Endosc* 18(5):790–795
5. Mack MJ (2001) Minimally invasive and robotic surgery. *J Am Med Assoc* 285(5):568–572
6. Hubens G, Coveliers H, Balliu L, Ruppert M, Vaneerdeweg W (2003) A performance study comparing manual and robotic assisted laparoscopic surgery using the da Vinci system. *Surg Endosc* 17(10):1595–1599
7. Judkins TN, Oleynikov D, Stergiou N (2005) Real-time augmented feedback benefits robotic laparoscopic training. *Stud Health Technol Inform* 119:243–248
8. Garcia-Ruiz A, Gagner M, Miller J, Steiner CP, Hahn JF (1998) Manual vs robotically assisted laparoscopic surgery in the performance of basic manipulation and suturing tasks. *Arch Surg* 133:957–961
9. DeUgarte DA, Etzioni DA, Gracia C, Atkinson JB (2003) Robotic surgery and resident training. *Surg Endosc* 17:960–963
10. Chang L, Satava RM, Pellegrini CA, Sinanan MN (2003) Robotic surgery: identifying the learning curve through objective measurement of skill. *Surg Endosc* 17(11):1744–1748
11. Narazaki K, Oleynikov D, Stergiou N (2006) Robotic surgery training and performance: identifying objective variables for quantifying the extent of proficiency. *Surg Endosc* 20(1):96–103
12. Ruurda JP, Broeders IAMJ (2003) Robot-assisted laparoscopic intestinal anastomosis. *Surg Endosc* 17(2):236–241
13. Lomanto D, Cheah W-K, So JB, Goh PM (2001) Robotically assisted laparoscopic cholecystectomy: A pilot study. *Arch Surg* 136(10):1106–1108
14. Judkins TN, Oleynikov D, and Stergiou N (2007) Objective evaluation of expert and novice performance during robotic surgical training tasks. *Surg Endosc* (in press)
15. Judkins TN, Oleynikov D, and Stergiou N (2007) Enhanced robotic surgical training using augmented visual feedback. *Surg Innov* (in press)
16. Jacobsen G, Elli F, Horgan S (2004) Robotic surgery update. *Surg Endosc* 18:1186–1191
17. Perez A, Zinner MJ, Ashley SW, Brooks DC, Whang EE (2003) What is the value of telerobotic technology in gastrointestinal surgery? *Surg Endosc* 17(5):811
18. D'Annibale AMD, Fiscun VMD, Trevisan PMD, Pozzobon MMD, Gianfreda VMD, Sovernigo GMD, Morpurgo EMD, Orsini CMD, Del Monte DMD (2004) The da vinci robot in right adrenalectomy: considerations on technique. *Surg Laparosc Endosc Percutan Tech* 14(1):38–41
19. Gray A (ed) (1997) *Modern differential geometry of curves and surfaces with mathematica*. CRC Press LLC, Boca Raton
20. Weisstein EW (2006) *Curvature*. [cited 2006 August]. Available from: <http://mathworld.wolfram.com/Curvature.html>
21. Campbell MJ, Gardner MJ (1988) Calculating confidence intervals for some non-parametric analyses. *Br Med J (Clin Res Ed)* 296(6634):1454–1456
22. Bates BT (1996) Single-subject methodology: an alternative approach. *Med Sci Sports Exerc* 28(5):631–638
23. Dufek JD, Bates BT (1991) Biomechanical factors associated with injury during landing in jump sports. *Sports Med* 12(5):326–337
24. Bates BT, James CR, Dufek JD (2004) Single-subject analysis. In: Stergiou N (ed) *Innovative analysis of human movement*. Human Kinetics, Champaign, IL, pp 3–28
25. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 236(4):458–464