SCIENTIFIC PAPER

Robotic Surgical Skills: Acquisition, Maintenance, and Degradation

Eric L. Jenison, MD, Karen M. Gil, PhD, Thomas S. Lendvay, MD, Michael S. Guy, MD

ABSTRACT

Background and Objectives: The degradation in robotic skills that occurs during periods of robotic surgical inactivity in newly trained surgeons was measured. The role of animate training in robotic skill was also assessed.

Methods: Robotically naïve resident and attending surgeons underwent training with the da Vinci® robot on needle passage (DN), rocking ring transfer peg board (RPB), and running suture pod tasks (SP). Errors were established to convert actual time to adjusted time. Participants were deemed "proficient" once their adjusted times were within 80% of those set by experienced surgeons through repeated trials. Participants did not use the robot except for repeating the tasks once at 4, 8, and 12 weeks (tests). Participants then underwent animate training and completed a final test within 7 days.

Results: Twenty-five attending and 29 resident surgeons enrolled; 3 withdrew. There were significant increases in time to complete each of the tasks, and in errors, by 4 weeks (Adjusted times: DN: 122.9 ± 2.2 to 204.2 ± 11.7, t=6.9, $P<.001$; RPB: 262.4 ± 2.5 to 364.7 ± 8.0, t=12.4, $P<.001$; SP: 91.4 ± 1.4 to 169.9 ± 6.8, t=11.3, $P<.001$). Times decreased following animate training, but not to levels observed after proficiency training for the RPB and SP modules.

Conclusions: Robotic surgical skills degrade significantly within 4 weeks of inactivity in newly trained surgeons. Animate training may provide different skills than those acquired in the dry lab.

Key Words: Robotic surgical skills, Training, Degradation.

Department of Obstetrics and Gynecology, Akron General Medical Center, Akron, Ohio, USA (Drs. Jenison, Gil, Guy).

Northeast Ohio Medical University, Rootstown, Ohio (Drs. Jenison, Gil).

Pediatric Urology, Seattle Children's Hospital, Seattle, Washington, USA (Dr. Lendvay).

Presentations: Jenison E, Gil K, Lendvay T, Guy M. Skill degradation and re-Training on the da Vinci robot. World Robotic Gynecology Congress II (WRGCII), Orlando, FL, April 11–13, 2010; Jenison E, Gil K, Lendvay T, Guy M. Skill acquisition and maintenance on the da Vinci® robot. World Robotic Gynecology Congress III (WRGCIII), Washington DC, May 5–6, 2011.; Gil KM, Guy M, Lendvay T, Jenison E. A curriculum for developing competency on the da Vinci® Robot. CREOG & APGO annual meeting, San Antonio, TX, March 9–12, 2011.

Funding: Skill Retention and Retraining Using the *da Vinci*® Robot, Eric Jenison MD (PI). Department of Defense, W81-XWH-08-1-0754. September 2008–March 2011.

Address correspondence to: Eric Jenison, MD, FACOG, Department of Obstetrics and Gynecology, Akron General Medical Center, 224 West Exchange St, #120, Akron OH 44302, USA. Telephone: (330) 996-2936, Fax: (330) 996-2912, E-mail: ejenison@agmc.org

DOI: 10.4293/108680812X13427982376185

INTRODUCTION

Since its introduction in 2001, the number of surgeons across multiple specialties adopting the *da Vinci* Surgical System has grown rapidly and yet little is understood about methodologies to develop surgeons who are skilled with this new technology. In the United States at the end of 2010, there were 1285 systems installed, and a total of 1752 installed worldwide.[1] That year, approximately 278 000 procedures were performed robotically, a 35% increase from 2009. Of these cases, the most commonly performed was the robotic total laparoscopic hysterectomy (110 000 cases, of which 32 000 were for malignancy). The robotic-assisted radical prostatectomy (RARP) was the second most common procedure with approximately 98 000 being performed worldwide.[1] At the end of 2008, between 75% and 85% of radical prostatectomies were completed robotically.[2]

This shift towards robotics has been driven by many factors including a significant reduction in morbidity.[3] These benefits stem from enhancements that include 3-dimensional visualization, 10X and zoom magnification, hand tremor dampening, and refined dexterity.[4] In gynecology, robotic-assisted radical hysterectomy (RARH) is associated with a decrease in blood loss, hospital stay, and similar operative times compared to laparoscopic and open procedures.[5] Robotics has also proven superior to other approaches in treatment of the obese and morbidly obese gynecologic population.[6] In the treatment of prostate cancer, RARP is associated with shorter hospitalization, fewer blood transfusions, comparable operative times, and a lower incidence of positive surgical margins compared to laparoscopic radical prostatectomy and open radical prostatectomy.[7]

Surgeons naïve to robotics form a heterogeneous group ranging from residents in training to experienced attendings. Currently, the optimal approach for training remains undefined, and the learning curve for proficiency is not well documented. Data suggest that formal training during fellowship is beneficial in improving surgical outcome in robotic prostatectomy.[8] How best to accomplish this training has yet to be determined. Animate training historically is an important modality in the surgical training process, and most current robotic credentialing protocols require a live porcine lab. This training is expensive and time consuming and, for some, an ethical issue. If this component is necessary for optimal training, then including it appears reasonable; however, the need for this component has not been explored.

As this technology continues to affect gynecologic, urologic, and ultimately other surgical fields, understanding the components of skills maintenance and degradation is paramount. Newly trained robotic surgeons may perform few robotic cases during the first months or year following training as this new technique is incorporated into their practice, and there may be significant degradation of skills during this time frame. Within traditional laparoscopy, mixed reports suggest skills degrade at 4 months[9] and 1 year,[10] while still others report no significant loss at 12 months.[11] Investigation of the extent to which skills on this complex system degrade during the early phases of use is important, because re-training may be beneficial. Determination of skill degradation over time would provide essential information for identification of potential re-training intervals to maintain surgical skills.

In this study, we seek to objectively document naïve surgeons' learning curves at the robotic console. Following this training, the natural degradation that occurs during periods of robotic surgical inactivity was monitored over 12 weeks. The role that animate training plays in robotic skill development was also assessed.

MATERIALS AND METHODS

Participants

This study was an IRB-approved prospective study funded through the Department of Defense. Resident and attending physicians in surgical subspecialties who had never received formal instruction on the *da Vinci* Surgical System were eligible to participate. All eligible surgeons who desired to participate during the study period were enrolled. This was a time-intensive study and attending physicians were compensated at a rate commensurate with their salary as surgeons ($250/hour). Resident physicians were given a $1000 educational stipend upon completion of the study. Attending surgeons were given the option of tuition reimbursement to attend the Intuitive animate training program or attend an animate training program developed by the investigators for this study (see below). Resident physicians were all trained at the investigator-developed animate program. Following a review of the study protocol and payment structure, informed consent was obtained.

Study Protocol

Enrolled participants completed a robotic surgery curriculum beginning with an on-line didactics-training module provided by Intuitive Surgical, Inc. (Sunnyvale, CA) that explained the basic principles of the robot. Participants then moved to the operating room (OR) where they completed a brief, introductory, inanimate training module using the *da Vinci* Surgical System, developed by Intuitive Surgical, Inc., under the supervision of the research assistant (RA). Subsequently, the 3 modules of the proficiency test (see below) were explained to the participants, including the adjusted times to complete the module that they needed to meet (actual time plus penalty time from errors). Participants engaged in repeated trials of each module until they met the adjusted time required (see description below). Once they had completed the proficiency test, participants did not use the robot for 12 weeks, except for repeating the 3 modules at 4, 8, and 12 weeks (tests). Following the 12-week test, participants underwent animate training at a lab developed by the investigators (see below) or at the Intuitive animate training lab. Participants completed a final test on the 3 modules within 7 days of completing the pig lab (PAT test).

Proficiency Test

The test consisted of 3 modules: a needle passage task (Dots/Numbers, DN) (http://www.thecgroup.com), a rocking ring transfer peg board task (RPB), and a running suture pod task (SP) **(Figure 1)**. These 3 tasks were chosen as representative skills tasks challenging practical abilities for robotic surgery. DN required physicians to pass a needle through 4 entrance and exit dots. RPB required a ring be picked up with one hand, passed to the other hand, and then placed on a specific peg while the board rocked back and forth. SP required a knot be tied (3

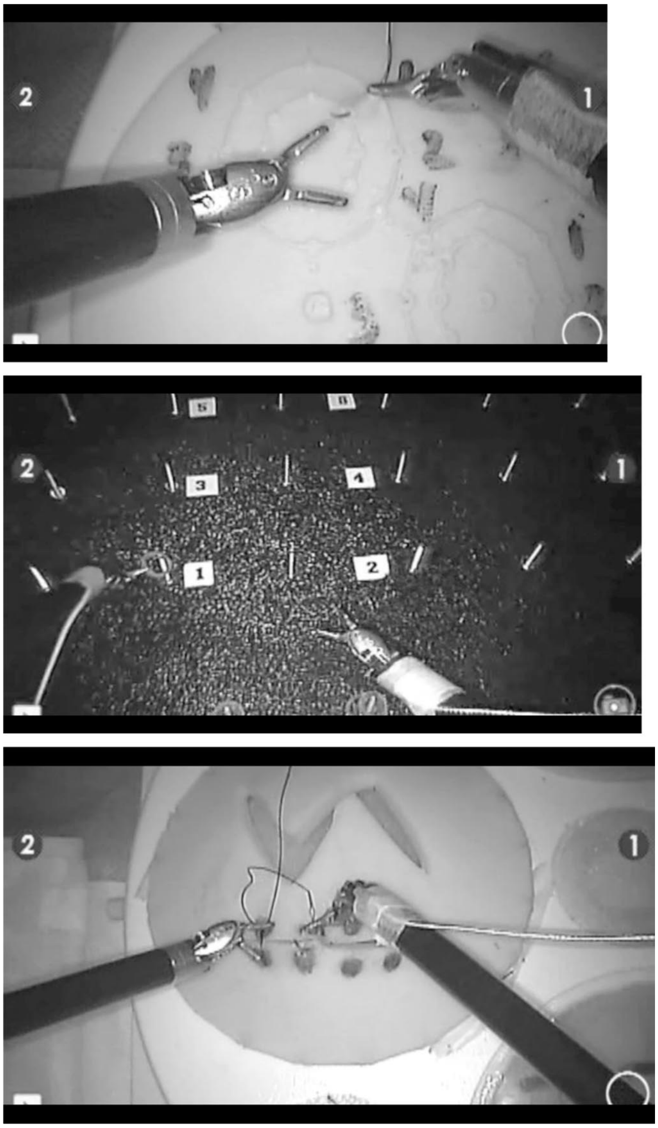

Figure 1. Modules of the proficiency test.

times) and then the needle be passed through 3 sets of dots. The DN and SP modules tested the targeting and passage of a suture needle as well as intracorporeal suturing, important for human extirpative and reconstructive surgery. The RPB required significant memory and concentration and was also chosen for its ability to discriminate targeting and precision, while challenging spatial relations capabilities. The benchmark performance metrics were established by experienced robotic surgeons. Three experienced surgeons completed each module until a plateauing of their task time was achieved. Technical and cognitive errors were established; a 5-second penalty/error rate was used to convert actual time to adjusted time (the ultimate performance metric). Participants were deemed "proficient" once they achieved performance to within 80% of the experienced surgeons' time through repeated trials. All trials were videotaped and scored by a trained rater.

Animate Training

A robotic pig lab curriculum was developed by one of the investigators (ELJ), which consisted of general robotic surgical techniques followed by general surgical, urologic, or gynecologic surgical tasks. The lab was held at a robotic training facility available at Ohio State University and was taught by physicians and the RA who had undergone training by one of the investigators (ELJ). All participants completed a cystotomy with a 2-layer closure, supercervical hysterectomy, cystectomy, and pelvic iliac artery and vein exposure; additional procedures included pelvic lymph node dissection (gynecology and urology), bladder reanastomosis at the urethra and bladder neck (urology), and colon resection (general surgery).

RESULTS

Twenty-five attending surgeons from general and vascular surgery, obstetrics/gynecology (ob/gyn), and podiatry were enrolled; 1 ob/gyn physician withdrew from the study leaving 24 attending surgeons **(Table 1)**. Twenty-nine resident surgeons from general surgery, urology, and ob/gyn were enrolled; 2 did not wish to continue and withdrew from the study (1 in general surgery and 1 in ob/gyn) leaving 27 resident surgeons **(Table 1)**. All resident physicians completed the animate training lab developed by the investigators. Eighteen of 24 attending physicians completed the animate training.

The proficiency test consisted of 3 modules (DN, RPB, and SP). The scoring algorithm for the DN module was changed after 10 resident physicians had initially completed the DN proficiency test, because a particular error was not recorded. These physicians were called back to repeat the test and achieve proficiency with the new algorithm; therefore, the number of trials to achieve proficiency could not be calculated for them. The SP module was added after 2 resident and 3 attending physicians started the study.

There was a wide range in number of trials required to achieve proficiency for each module **(Figure 2, Table 2)**. Twenty-three of 24 attending surgeons achieved proficiency on the DN module within 8 trials, 1 surgeon

Table 1.
Characteristics of Resident and Attending Physicians

	Resident Physicians (n=27)	Attending Physicians (n=24)
Specialty (N)		
General Surgery	5	4
Ob/Gyn	14	15
Podiatry		2
Urology	8	
Vascular Surgery		3
Age (Mean ± SEM, range)	29.9 ± 0.6 (26–38)	41.4 ± 1.7 (30–62)
Male (%)	70	71
PGYear (Residents)		
1	3	
2	7	
3	6	
4	8	
5	2	
6	1	
Self-reported number of laparoscopic cases assisted (residents) or performed (attendings) last year; (Mean ± SEM, range)	41.7 ± 6.8 (0–150)	85.2 ±17.5 (0–350)
Right Handed (%)	96	83
Ever played musical instrument on a regular basis (more than 3 times/week) (%Yes)	52	67
Ever played video games (%Yes)	100	75
Total years played (Mean ± SEM, range)	16.4 ± 1.3 (1–30)	9.2 ± 2.1[a] (0–31)

[a]t=2.956, P=.005, Attending versus Resident Physicians. Standard error of the mean (SEM).

required 20 trials. One attending surgeon required 41 trials to achieve proficiency on the RPB and SP, and another attending surgeon required 39 trials to achieve proficiency on the SP. Mean number of trials to reach proficiency on the RPB module was significantly lower for resident physicians **(Table 2)**, regardless of whether the score from the one participant who required 41 trials was included. Despite requiring more trials to achieve proficiency, data from this participant did not differ significantly from that of other participants, and there were no differences in further analyses of data with and without this participant; thus, the participant was included in the analyses.

Due to scheduling conflicts, a range existed in the number of days between the last trial of the proficiency test and the 4-week test **(Table 3)**. There was a wide range in time required to complete each of the modules on the 4-week tests **(Table 3)**. For all modules, a significant increase occurred in the mean adjusted time to complete the modules at the 4-week test relative to the last trial on the proficiency test (all $P < .001$, **Table 3**).

No significant differences were found in mean adjusted times on the 4-week test between resident and attending surgeons for any of the modules. Data from resident and attending surgeons were therefore combined. Mean adjusted times at 4 weeks did not differ as a function of sex

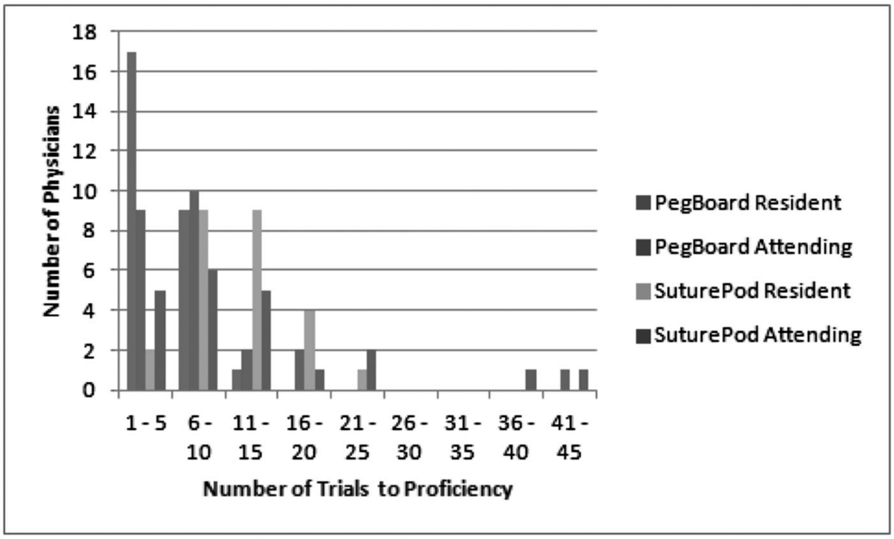

Figure 2. Number of trials required to achieve proficiency on RPB and SP.

Table 2.
Number of Trials to Achieve Passing Score for the Proficiency Test

Modules	Resident Physicians (n=27)	Attending Physicians (n=24)
DN	4.2 ± 0.5 (1–9)[a]	4.6 ± 0.8 (2–20)
RPB	5.2 ± 0.4 (2–11)	8.8 ± 1.7 (2–41)[c]
SP	11.5 ± 0.9 (5–23)[b]	13.0 ± 2.3 (1–41) [b]

Mean ± SEM (range).

[a]n=17 as 10 residents were called back to repeat the test using a modified scoring algorithm.

[b]n=25 for residents physicians; n=21 for attending physicians for SP module.

[c]t=2.088, P=.047, Attending versus Resident Physicians.

or history of playing an instrument for any of the modules. No significant correlation existed between time to complete any of the modules at 4 weeks and age or self-reported number of laparoscopic cases performed last year. An inverse relationship existed between self-reported number of years playing video games and time to complete the DN module (r=−.298, P=.034) and a positive relationship with the number of days between the last trial of the DN proficiency test and the 4-week DN test (r=.303, P=.031). These correlations were not significant for the other 2 modules.

Mean number of days between the 4- and 8-week tests was 31.5 ± 1.1 (range, 14 to 56); between the 8- and 12-week tests was 32.6 ± 1.2 (range, 24 to 70); and between the 12-week and PAT test was 45.3 ± 4.5 (range, 2 to 152). The PAT was conducted within 7 days of the animate training lab.

No significant differences existed in the adjusted time to complete the 3 modules between attending and resident physicians during the 4-, 8-, 12-, and PAT tests (all P>.05) and no significant interactions between time and resident or attendings (all P>.05); therefore, data from resident and attending physicians were analyzed together **(Figures 3, 4, and 5)**. Actual time increased significantly between the last trial of the proficiency test and the 4-week test (DN: 106.9 ± 2.0 to 161.5 ± 8.6, t=6.4, P<.001; RPB: 233.0 ± 3.3 to 305.2 ± 7.9, t=9.2, P<.001; SP: 87.3 ± 1.5 to 158.8 ± 6.2, t=11.5, P<.001). The main error made on the DN module was missing the dot at entry or exit (mean number of errors increased from 3.1 ± 0.2 to 8.2 ± 0.8, t= 6.4, P<.001); there was also an increase in number of needle drops (mean number of errors increased from 0.06 ± 0.03 to .3 ± 0.07, t=2.7, P=.01). The main errors made on the RPB module were instruments out of view and instrument touching the peg; these errors increased significantly (mean number of instruments out of view increased from 0.37 ± 0.08 to 1.2 ± 0.2, t=3.5, P=.001; mean number of instruments touching the peg increased from 4.5 ± 0.3 to 9.3 ± 0.5, t=11.4, P<.001). The main error made on the SP module was missing the dot at entry or exit (mean number of errors increased from 0.65 ± 0.09 to 1.9 ± 0.3, t=4.4, P<.001). Adjusted time was significantly increased at the 4-week test relative to the last trial of the proficiency test (DN: 122.9 ± 2.2 to 204.2 ± 11.7, t=6.9, P<.001; RPB: 262.4 ± 2.5 to 364.7 ± 8.0, t=12.4, P<.001; SP: 91.4 ± 1.4 to 169.9 ± 6.8, t=11.3, P<.001).

Table 3.
Adjusted Time on Last Trial of Proficiency Test (PT), Number of Days Between PT and 4-Week Test, and Adjusted Time on 4-Week Test

	Resident Physicians (n=27)	Attending Physicians (n=24)
Days between last trial of DN PT and 4-week DN test	36.7 ± 3.6 (19–104)	40.1 ± 3.3 (21–72)
Adjusted time on last trial of DN PT	124.6 ± 2.7 (97–150)	121.0 ± 3.6 (75–143)
Adjusted time on 4-week DN test[a]	194.7 ±15.1 (86–417)	215.0 ± 18.2 (100–375)
Days between last trial of RPB PT and 4-week RPB test	42.5 ± 3.3 (19–84)	33.8 ± 2.4 (21–70)[c]
Adjusted time on last trial of RPB PT	265.1 ± 3.0 (229–285)	259.4 ± 4.2 (206–297)
Adjusted time on 4-week RPB test[a]	360.0 ± 11.6 (259–484)	370.3 ± 11.0 (279–456)
Days between last trial of SP PT and 4 week SP test	31.2 ± 1.4 (22–55)	30.9 ± 1.2 (21–46)
Adjusted time on last trial of SP PT[b]	93.9 ± 2.4 (82–129)	88.5 ± 1.2 (75–95)
Adjusted time on 4 week SP test[a]	177.9 ± 8.0 (124–287)	160.3 ± 11.1 (81–282)

Time (Seconds), Mean ± SEM (range).

[a]Paired t-test, 4 week test versus last trial of PT, all P<.001 (DN: residents, t=4.9, attendings, t=4.9; RPB: residents, t=7.9, attendings, t=9.8; SP: residents, t=9.6, attendings, t=6.5).

[b]n=25 for Residents Physicians; n=21 for Attending Physicians for SP module.

[c]t=2.079, P=.043, Attending versus Resident Physicians.

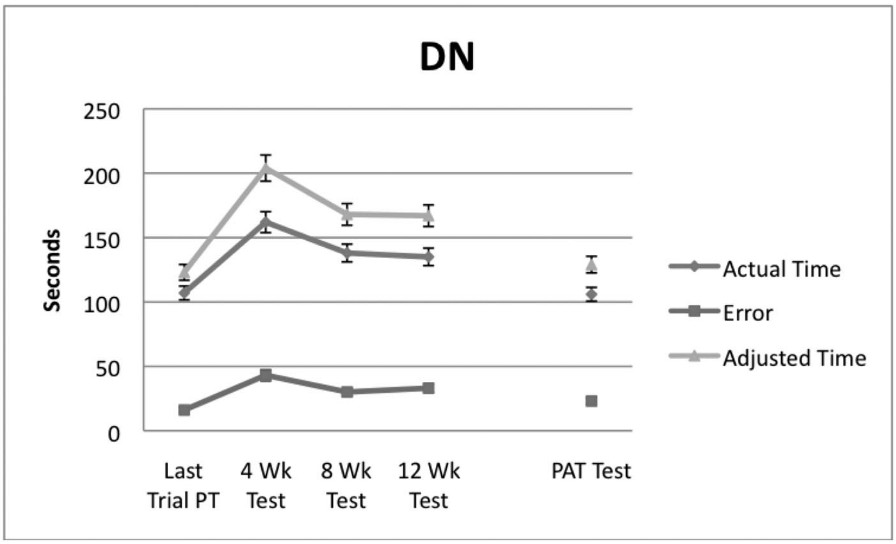

Figure 3. Actual time, penalty time for errors (each error = 5 seconds), and adjusted time over the 12 weeks of the study for resident and attending physicians completing the DN module (n=51). Actual times and penalty time for errors were significantly different over the 12 weeks of the study (Actual time, F=15.765, P<.001; Errors, F=17.072, P<.001;Adjusted time, F=18.358, P<.001). Mean ± SEM.

Actual times for the DN module decreased between the 4- and 8-week tests and remained constant at the 12-week test **(Figure 3)**. Actual times on the RPB and SP remained fairly constant over the 4-, 8-, and 12-week test **(Figures 4 and 5)**. Error rates for all modules were fairly constant over the 4-, 8-, and 12-week tests **(Figures 3, 4, and 5)**.

Actual times decreased significantly between the 12-week and PAT tests on the DN and RPB modules (DN, t=3.9, P<.001; RPB, t=3.6, P=.001) **(Figures 3 and 4)**. Errors decreased for the DN and RPB module after the PAT test (DN, t=3.1, P=.004; RPB, t=5.4, P<.001). As a result, actual and adjusted times on the DN module were not significantly different from those obtained on the last trial

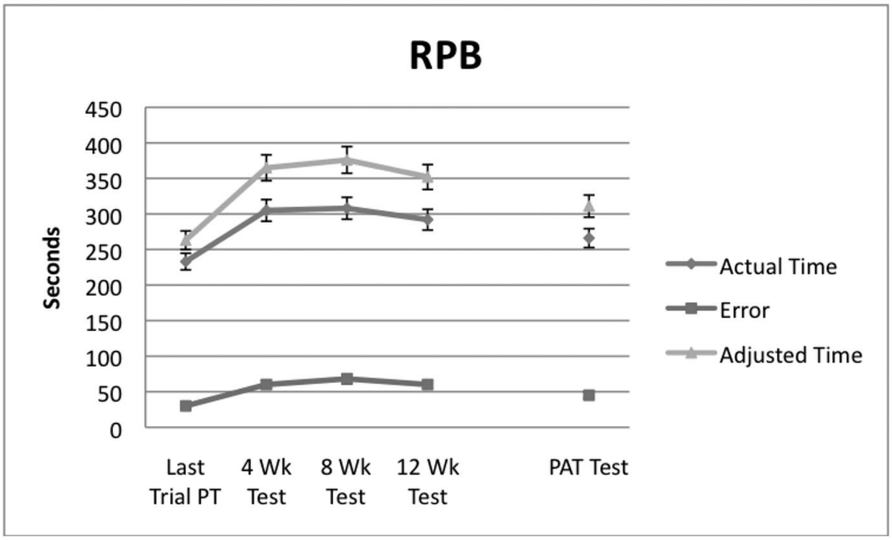

Figure 4. Actual time, penalty time for errors (each error = 5 seconds), and adjusted time over the 12 weeks of the study for resident and attending physicians completing the RPB module (n=51). Actual times and penalty time for errors were significantly different over the 12 weeks of the study (Actual time, F=43.528, P<.001; Errors, F=63.366, P<.001; Adjusted time, F=74.891, P<.001). Mean ± SEM.

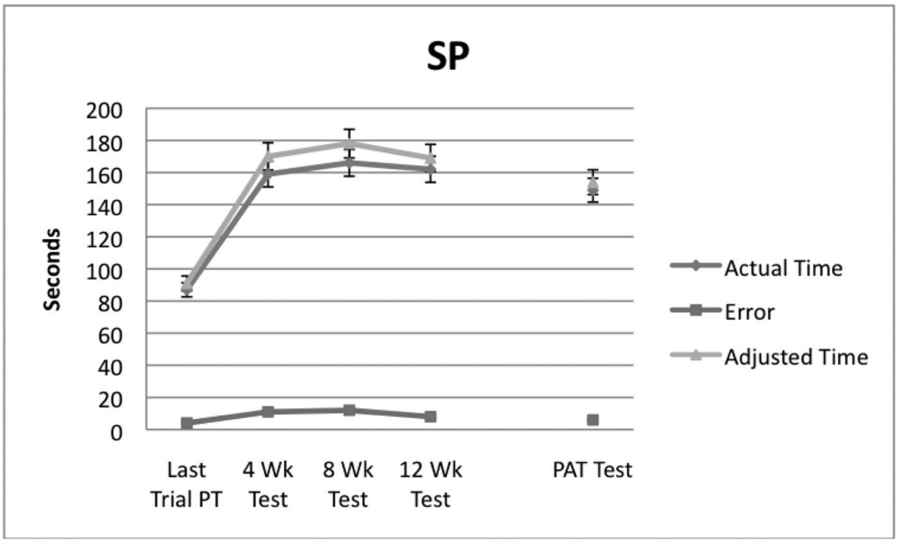

Figure 5. Actual time, penalty time for errors (each error = 5 seconds), and adjusted time over the 12 weeks of the study for resident and attending physicians completing the SP module (n=46). Actual times and penalty time for errors were significantly different over the 12 weeks of the study (Actual time, F=55.502, P<.001; Errors, F=8.840, P<.001; Adjusted time, F=51.556, P<.001). Mean ± SEM.

of the proficiency test, although errors were still higher (t=2.9, P=.006). For the RPB, actual and adjusted times on the PAT test, as well as errors, were greater after the PAT than they were on the last trial of the proficiency test (actual time, t=4.9, P<.001; errors, t=5.6, P<.001; adjusted time, t=5.9, P<.001). For the SP module, actual and adjusted times, but not errors, were significantly greater after the PAT than they were during the last trial of the proficiency test (**Figure 5**, actual times, t=9.6, P<.001; adjusted time, t=9.5, P<.001).

Assessment was made of the degree to which an individual's time to complete the modules was correlated with subsequent times at each re-testing interval. Beginning with the 4-week test, time to complete RPB and SP at one time point was correlated with time to complete it on the subsequent time point **(Table 4)**.

The degree to which an individual's time to complete one module was correlated with time to complete the other modules was assessed **(Table 5)**. At 8 weeks, there was a significant correlation between time to complete the RPB and SP modules; by 12 weeks the correlations were significant for RPB and SP and for DN and SP. During the PAT test, time to complete one module was correlated with time to complete the other modules, with the correlation between RPB and SP reaching .646 ($P<.001$).

DISCUSSION

Robotic skills degraded significantly within 4 weeks of surgical robotic inactivity. The degradation was seen with the simpler DN module as well as with the more complex RPB module, which involved camera and arm clutching as well as arm movements, and SP module, which involved intracorporeal knot tying. Analysis of errors during the tests also sheds light on skill level. Between the last proficiency trial and the 4-week test, the number of errors made in all modules increased significantly and remained elevated at 8 weeks and 12 weeks. Even after animate training, the error rates remained significantly higher than following the initial proficiency test for DN and RPB. This may carry important safety implications for new robotic surgeons who have gone though the required training, but have had several weeks lapse without time on the console. These findings may also be germane to hospital credentialing, annual re-credentialing and Maintenance of Certification (MOC) boards responsible for ensuring that their surgeons remain proficient.

At 8 weeks, performance improved slightly on the DN test, and remained stable on the RPB and SP tests. It may be that the testing situations at 4, 8, and 12 weeks served as mini-training sessions, so that average performance on the later tests did not continue to decline. Additionally, the DN module may have served as a "warm up" trial for the other modules as this module was always conducted first. As the study progressed, performances on the tests began to correlate so that surgeons who did well on the modules at one time point also did well on the modules at later time points. Thus, surgeons who performed well at 4 weeks were likely to perform well at later time points. Conversely, surgeons who took longer to complete modules at 4 weeks were likely to require more time to complete modules at later time points. These variations demonstrate unique individual learning curve patterns and suggest some surgeons may require more intense or more frequent re-training following their initial training.

Performance among modules began to correlate over the study so that a surgeon's performance on one module became increasingly correlated with performance on other modules. By the 12-week test, surgeons who performed well on the SP test were also performing well on the RPB and DN tests, and by the PAT test, surgeons' performances on all modules were significantly correlated. By the PAT test, surgeons who did not perform well on one module did not perform well on the other modules. This suggests that some surgeons were becoming increasingly versatile with the robot, able to perform equally well on a variety of tasks while others did not perform as well on any of the tasks.

Individual variations were apparent from the beginning of the study. There was a large variation in the number of attempts required to achieve proficiency. This was true for both attending and resident surgeons. There was no difference in the mean number of attempts to achieve proficiency between resident and attending physicians with the exception of RPB, in which the residents achieved proficiency significantly faster. These objective data support observational data from the literature demonstrating a wide range in the number of cases required to achieve proficiency in urology and gynecology.[12–15]

Table 4.
Correlations Between Adjusted Time on One Test and the Subsequent Test

	Last Trial of PT and 4-Week Test	4-Week Test and 8-Week Test	8-Week Test and 12-Week Test	12-Week Test and PAT Test
DN	.022	.227	.192	.289
RPB	.053	.326^a	.315^a	.440^b
SP	−.013	.554^c	.310^a	.348^a

Values are Pearson Correlation Coefficients. $^a P<.05$; $^b P<.01$; $^c P<.001$.

Table 5.
Correlations Among Adjusted Time to Complete the Modules over the Study

	Last Trial of PT	4-Week Test	8-Week Test	12-Week Test	PAT Test
DN and RPB	.120	.178	−0.069	.249	.331[a]
RPB and SP	.137	.242	.299[a]	.364[a]	.646[c]
DN and SP	−.287	.161	.023	.332[a]	.410[b]

Values are Pearson Correlation Coefficients. [a]$p<.05$; [b]$p<.01$; [c]$p<.001$.

In 1936, Wright introduced the concept of the "learning curve," proposing a mathematical model for the aircraft industry.[16] Working definitions of the learning curve in surgery have varied widely. These include the diminishing amount of time to perform a specific task,[17] or conversely, a self-declared point at which a surgeon reaches a comfort zone when performing a procedure.[17] Although, no standard definition has been accepted,[18] it is often defined as the number of cases a surgeon needs to perform to achieve acceptable operative times and reasonable outcomes.[19]

The most efficient method for mastering a skill is poorly understood. For the last century, surgeons have been trained in the method promoted by Halsted, in which the surgeon is gradually exposed to increasingly difficult tasks until they are able to operate autonomously.[20–22] Under these methods, there is wide variation; the estimated learning curve in RARP ranges from 25[13] to 150 cases.[14] Seamon et al[15] report that 20 cases are required to achieve proficiency in RARH.

Previous studies point to a correlation between performing a high volume of complex cases and a decrease in operative mortality.[23–26] It follows conversely that the patients cared for by a low-volume surgeon have an increased relative risk of mortality in the same procedures.[27] Traditionally, being privileged for a procedure implied that a surgeon would be able to perform this operation indefinitely, regardless of actual practice. However, a Report of a Discussion and Study Group of the American Surgical Association recognizes that re-privileging should be based on verifiable criteria and linked to practice outcomes.[28] These endpoints of assessment are not clearly determined yet, but the implication is that without sufficient exposure, surgical skills degrade.

Given the differences in the extent of performance degradation at 4 weeks, analyses were performed to assess whether there are variables that can be used to predict skill degradation and/or maintenance. There were no significant differences in times between resident and attending physicians at 4 weeks. Additional factors that were examined were sex, age, history of playing a musical instrument, self-reported number of laparoscopic surgeries performed the last year or number of years playing video games. The only significant correlation was between the self-reported years playing video games and time to complete the DN module at 4 weeks. These data suggest that observation and testing on an individual basis may be required to assess the proficiency of robotic surgeons.

Robotic training currently includes an animate training lab. This lab covers surgical specialty specific procedures and provides experience with handling of tissue. For the simpler skill (DN), animate training restored performance to levels achieved after repeated trials of the inanimate proficiency test; however, skill level on the RBP and SP did not return to the levels achieved on the inanimate test. This suggests that different skills are being learned in the 2 settings and that animate training did not provide the kind of intense and specific training on how to manipulate the robot that inanimate training did. The *da Vinci* Surgical System is a complex system, and it may be that becoming skilled with this system requires more inanimate training than is currently prescribed.

This may also highlight the differences in the core components that make a robotic surgeon. Mechanical fundamentals include both knowledge of how the console is operated and manipulation skills that are required to operate the robot as a surgical tool to efficiently perform tasks without error. The second component involves anatomic fundamentals learned during medical school and cemented by repetition in the operating room. This includes knowledge of anatomy, the handling of various tissues and dissection planes and the skills to navigate spaces to achieve a desired surgical outcome. The porcine lab may emphasize the latter, allowing surgeons to translate their haptic and visual feedback learned in open and laparoscopic cases to the robotic console. While the porcine lab did not bring surgeons back to objectively measured inanimate proficiency levels, it is possible that more subjective, yet fundamental, aspects of robotic surgery were impacted. Future research will explore skills acquired during live animate training and whether these skills can be acquired in less expensive and less time-consuming inanimate settings.

Limitations of this study are that it is a single institution study, conducted in a fairly homogenous group of surgical residents and attendings. Replication of these results in more diverse training programs and with surgeons with more varied backgrounds, skill levels, and experience is warranted. The fact that there were no significant differences between the means of residents and attendings at any of the test periods (4-, 8-, 12-week and PAT) suggests, however, that these results are fairly robust.

These results demonstrate that robotic surgical skills degrade rapidly in newly trained surgeons, and efforts need to be made to provide robotic surgeons with active curricula aimed at maintaining performance during periods of inactivity to ensure patient safety. Future research will focus on mechanisms to ensure that robotic skills are upheld in newly trained surgeons until they are routinely and consistently performing robotic surgery. Future research defining the training interval to maintain proficiency may start with re-training at 4-week intervals. The possibility of using virtual reality robotic surgery platforms as an adjunct for this maintenance of skills will be explored.

References:

1. Intuitive Surgical Inc. Annual Report 2010. Sunnyvale, CA. 1 Feb 2011. <http://phx.corporate-ir.net/phoenix.zhtml?c=122359&p=irol-sec>.

2. Lepor H. Status of radical prostatectomy in 2009: is there medical evidence to justify the robotic approach? *Rev Urol.* 2009 Spring;11(2):61–70.

3. Paley PJ, Veljovich DS, Shah CA, et al. Surgical outcomes in gynecologic oncology in the era of robotics: analysis of first 1000 cases. *Am J Obstet Gynecol.* 2011;204(6):551.e. 1–9. Epub 2011 Mar 16.

4. Chen CC, Falcone T. Robotic gynecologic surgery: past, present, and future. *Clin Obstet Gynecol.* 2009 Sep;52(3):335–343.

5. Magrina JF, Kho RM, Weaver AL, Montero RP, Magtibay PM. Robotic radical hysterectomy: comparison with laparoscopy and laparotomy. *Gynecol Oncol.* 2008 Apr;109(1):86–91.

6. Gehrig PA, Cantrell LA, Shafer A, Abaid LN, Mendivil A, Boggess JF. What is the optimal minimally invasive surgical procedure for endometrial cancer staging in the obese and morbidly obese woman? *Gynecol Oncol.* 2008 Oct;111(1):41–45.

7. Berryhill R Jr., Jhaveri J, Yadav R, et al. Robotic prostatectomy: a review of outcomes compared with laparoscopic and open approaches. *Urology.* 2008 Jul;72(1):15–23.

8. Kwon EO, Bautista TC, Jung H, Goharderakhshan RZ, Williams SG, Chien GW. Impact of robotic training on surgical and pathologic outcomes during robot-assisted laparoscopic radical prostatectomy. *Urology.* 2010 Aug;76(2):363–368.

9. Kahol K, Ashby A, Smith M, Ferrara JJ. Quantitative evaluation of retention of surgical skills learned in simulation. *J Surg Educ.* 2010 Nov-Dec;67(6):421–426.

10. Hiemstra E, Kolkman W, van de Put MA, Jansen FW. Retention of basic laparoscopic skills after a structured training program. *Gynecol Surg.* 2009 Sep;6(3):229–235.

11. Rosenthal ME, Ritter EM, Goova MT, et al. Proficiency-based fundamentals of laparoscopic surgery skills training results in durable performance improvement and a uniform certification pass rate. *Surg Endosc.* 2010 Oct;24(10):2453–2457.

12. Patel VR, Tully AS, Holmes R, et al. Robotic radical prostatectomy in the community setting–the learning curve and beyond: initial 200 cases. *J Urol.* 2005;174:269.

13. Herrell SD, Smith JA Jr. Robotic-assisted laparoscopic prostatectomy: what is the learning curve? *Urology.* 2005;66:105.

14. Guillonneau B, Rozet F, Barret E, Cathelineau X, Vallancien G. Laparoscopic radical prostatectomy: assessment after 240 procedures. *Urol Clin North Am.* 2001;28(1):189–202.

15. Seamon LG, Fowler JM, Richardson DL, et al. A detailed analysis of the learning curve: robotic hysterectomy and pelvic-aortic lymphadenectomy for endometrial cancer. *Gynecol Oncol.* 2009 Aug;114(2):162–167.

16. Wright TP. Factors affecting the cost of airplanes. *J Aeronaut Sci.* 1936;3:122.

17. Steinberg PL, Merguerian PA, Bihrle W 3rd, Seigne JD. The cost of learning robotic-assisted prostatectomy. *Urology.* 2008 Nov;72(5):1068–1072.

18. Herrell SD, Smith JA Jr. Robotic-assisted laparoscopic prostatectomy: what is the learning curve? *Urology.* 2005 Nov;66(5 Suppl):105–107.

19. Artibani W, Novara G. Cancer-related outcome and learning curve in retropubic radical prostatectomy: "if you need an operation, the most important step is to choose the right surgeon." *Eur Urol.* 2008 May;53(5):874–876.

20. Halsted WS. The training of the surgeon. *Am Med.* 1904;9:69–75.

21. Heuer GW. Dr. Halsted. *Bull Johns Hopkins Hosp.* 1952;90(2Suppl):1–105.

22. Osborne MP. William Stewart Halsted: his life and contributions to surgery. *Lancet Oncol.* 2007 Mar;8(3):256–265.

23. Hannan EL, Radzyner M, Rubin D, Dougherty J, Brennan MF. The influence of hospital and surgeon volume on in-hospital mortality for colectomy, gastrectomy, and lung lobectomy in patients with cancer. *Surgery.* 2002;131:6–15.

24. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA.* 1998;280:1747–1751.

25. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high volume hospitals: estimating potentially avoidable deaths. *JAMA.* 2000;283:1159–1166.

26. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med.* 2002;137:511–520.

27. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med.* 2003 Nov 27;349(22):2117–2127.

28. Bass BL, Polk HC, Jones RS, et al. Surgical privileging and credentialing: a report of a discussion and study group of the American Surgical Association. *J Am Coll Surg.* 2009 Sept;209(3): 396–404.