

Auditory Force Feedback Substitution Improves Surgical Precision during Simulated Ophthalmic Surgery

Nathan Cutler,¹ Marcin Balicki,² Mark Finkelstein,² Jiangxia Wang,³ Peter Geblbach,¹ John McGready,³ Iulian Iordachita,² Russell Taylor,² and James T. Handa¹

PURPOSE. To determine the extent that auditory force feedback (AFF) substitution improves performance during a simulated ophthalmic peeling procedure.

METHODS. A 25-gauge force-sensing microforceps was linked to two AFF modes. The “alarm” AFF mode sounded when the force reached 9 mN. The “warning” AFF mode made beeps with a frequency proportional to the generated force. Participants with different surgical experience levels were asked to peel a series of bandage strips off a platform as quickly as possible without exceeding 9 mN of force. In study arm A, participants peeled with alarm and warning AFF modes, the order randomized within the experience level. In study arm B, participants first peeled without AFF, then alarm or warning AFF (order randomized within the experience level), and finally without AFF.

RESULTS. Of the 28 “surgeon” participants, AFF improved membrane peeling performance, reducing average force generated ($P < 0.01$), SD of forces ($P < 0.05$), and force \times time above 9 mN ($P < 0.01$). Short training periods with AFF improved subsequent peeling performance when AFF was turned off, with reductions in average force, SD of force, maximum force, time spent above 9 mN, and force \times time above 9 mN (all $P < 0.001$). Except for maximum force, peeling with AFF reduced all force parameters ($P < 0.05$) more than peeling without AFF after completing a training session.

CONCLUSIONS. AFF enables the surgeon to reduce the forces generated with improved precision during phantom membrane peeling, regardless of surgical experience. New force-sensing surgical tools combined with AFF offer the potential to

enhance surgical training and improve surgical performance. (*Invest Ophthalmol Vis Sci.* 2013;54:1316-1324) DOI: 10.1167/iovs.12-11136

Vitreoretinal microsurgery involves the manipulation of delicate tissues and structures that can be unforgiving. These manual manipulations require high levels of dexterity and visual-motor coordination that are often at the limits of human capability. Due to the micron-level size of structures being manipulated, tactile sensation, referred to as haptic or force feedback, is often diminished or even absent. We recently measured tool-to-tissue forces during vitreoretinal surgery on rabbit eyes and found that the force necessary to induce a retinal tear was <7 mN.¹ At this level of force, surgeons can feel only 19% of the events performed.² These results indicate that the majority of vitreoretinal surgery is performed solely with visual feedback, which has been shown to reduce speed and accuracy of manipulations.³

Although it seems intuitive that adding force feedback to surgical instruments would provide measurable benefits and improve outcomes, the results are not unanimous.⁴ Benefits of force feedback were least disputed when added to robotic surgery, a type of surgery that has no force feedback in its current form. In addition to adding haptic force feedback, visual and auditory force feedback substitution can improve consistency of applied forces during maneuvers such as robotically assisted knot tying and in the acquisition of advanced skills for surgical trainees.⁵⁻⁹

In contrast to the incorporation of force feedback into minimally invasive surgery, where the goal is to regain lost senses, the addition of force-sensing into ophthalmic surgery, like robotic surgery, adds force information that was never before present. A common procedure for vitreoretinal surgeons is to peel thin membranes of scar tissue from the underlying retina. Peeling these epiretinal membranes is technically challenging because the surgeon must rely on nearly imperceptible visual cues to avoid undesirable forces between the surgical instrument and the tissue. By incorporating force-sensing elements into the tip of the instrument, vitreoretinal surgical tools have been created to accurately give information about forces applied during surgery.¹⁰ Our group has developed an auditory force feedback (AFF) system combined with a force-sensing instrument that is capable of alerting the surgeon of the applied force during a maneuver.^{1,10,11} The impact of this system on a surgeon's technical performance is unknown. We hypothesize that the addition of AFF would allow quantifiably better control and precision of applied forces during a simulated epiretinal membrane peeling task. To address this hypothesis, we conducted a prospective surgical trial that tested simulated surgical performance with and without auditory force feedback, and also determined the effects of training with

From the ¹Wilmer Eye Institute and the ²Engineering Research Center for Computer-Integrated Surgical Systems and Technology, Johns Hopkins School of Medicine, Baltimore, Maryland; and the ³Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland.

Supported in part by the U.S. National Science Foundation under Cooperative Agreement EEC9731478, the National Institute of Biomedical Imaging and Bioengineering/National Institutes of Health (NIH) Grant BRP 1 R01 EB 007969, the Achievement Rewards for College Scientists Foundation, Research to Prevent Blindness, Johns Hopkins internal funds, and the Wilmer Biostatistics/NIH Core Grant EY01765. Partial equipment support was provided by Alcon and Carl Zeiss Meditec. The authors alone are responsible for the content and writing of the article.

Submitted for publication October 13, 2012; revised December 13, 2012; accepted January 10, 2013.

Disclosure: N. Cutler, None; M. Balicki, None; M. Finkelstein, None; J. Wang, None; P. Geblbach, None; J. McGready, None; I. Iordachita, None; R. Taylor, None; J.T. Handa, None

Corresponding author: James T. Handa, Wilmer Eye Institute, Johns Hopkins School of Medicine, 400 N. Broadway, Smith Building Room 3015; Baltimore, MD 21287-9277; jthanda@jhmi.edu.

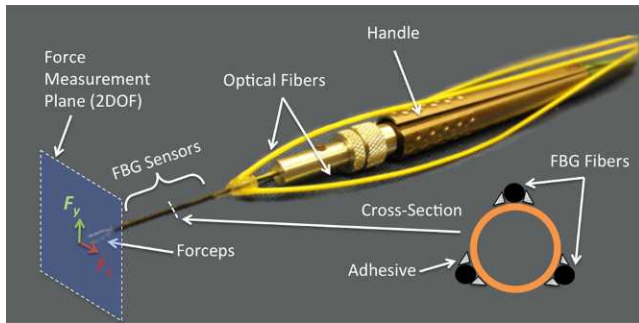


FIGURE 1. Diagram of the force sensing microforceps. Three FBG fiber sensors have been glued to the instrument shaft. The forceps tips detect forces in the x - and y -axes (2 degree of freedom plane).

AFF and whether the effects of AFF differed between novice and experienced retinal surgeons.

METHODS

Human Subjects

Twenty-eight participants, or “surgeons,” were enrolled in this study after obtaining written, informed consent. The group was comprised of 10 medical students, nine ophthalmology residents (three postgraduate year II, three postgraduate year III, three postgraduate year IV), five retina fellows, and four attending retina surgeons at The Johns Hopkins Hospital, Baltimore, MD. This study was approved by the Johns Hopkins University Institutional Review Board. The research followed the tenets of the Declaration of Helsinki.

Force-Sensing Tool and Recording System

For this study, reusable 25-gauge microforceps (Alcon, Inc., Fort Worth, TX) were customized to include three fiber Bragg grating (FBG) sensors bonded to the surface of the tool shaft, as described in detail by He et al.¹² FBGs are optical fiber sensors capable of detecting fine changes in strain, are immune to electrical interference, and can be sterilized. The

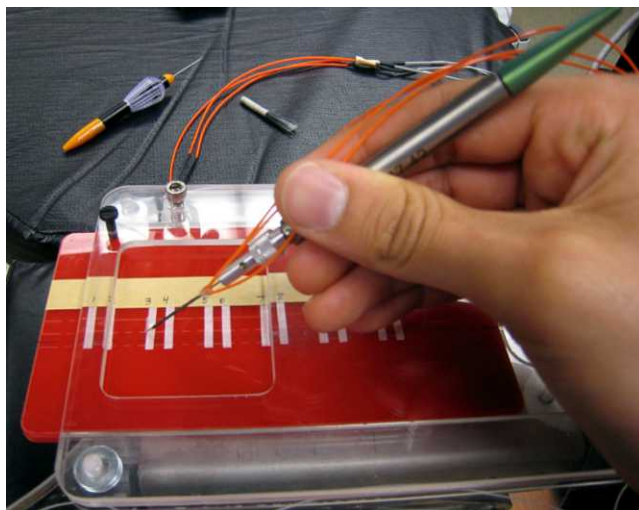


FIGURE 2. Experimental setup of phantom epiretinal membrane peeling apparatus with the fiber-optic force-sensing microforceps. The microforceps were connected to a computer that provided AFF. Participants were instructed to grab the top edge of the phantom membrane and peel toward themselves as quickly as they could without exceeding 9 mN of force.

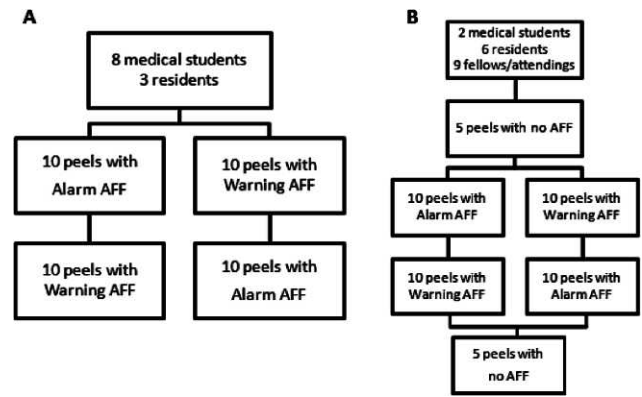


FIGURE 3. Study design for (A) the first 11 participants and (B) the last 17 participants.

deflection of the tool, caused by forces at the tip, generates axial strain of the FBGs, which is measured with an optical sensing interrogator (Model sm 130-700; Micron Optics, Inc., Atlanta, GA). The instrument reads forces applied at the tip in the x - and y -planes perpendicular to the instrument shaft (2 degrees of freedom [2DOF]), that is, force vector components applied along the main axis of the instrument are not measured, as shown in Figure 1. The overall sensitivity is 0.25 mN in the range of 0 to 60 mN.

Force data were acquired at a rate of 2000 Hz through a custom utility software package (Multi-Input Data Collection Utility, built using the CISST framework, a collection of libraries for development of computer-assisted intervention systems; InfinityQS, Fairfax, VA).¹³ A commercial computer (Dell Precision 690 workstation [Dell, Inc., Round Rock, TX], with a Xeon quad core processor and 2GB of random access memory running Linux operating system, Ubuntu 10.04 [Canonical Group Ltd., London, UK]). Video was recorded through the microscope in mono at 30 frames/s with a 1/3-inch camera (CCD IEEE1394b, Point Grey Flea2, model FL2-08S2C-C, 1032

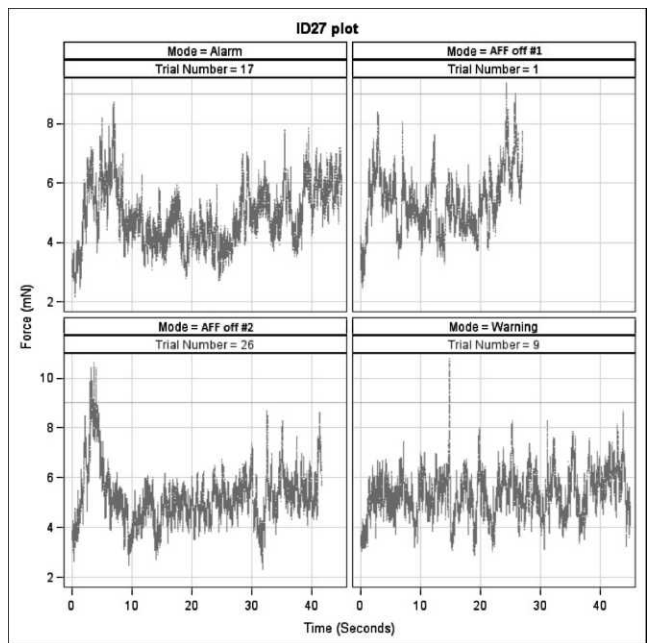


FIGURE 4. Representative force plots of four different membrane peels performed by one participant (ID27) with alarm AFF, AFF off #1, AFF off #2, and warning AFF. Trial number indicates the peeling order from 1 to 30.

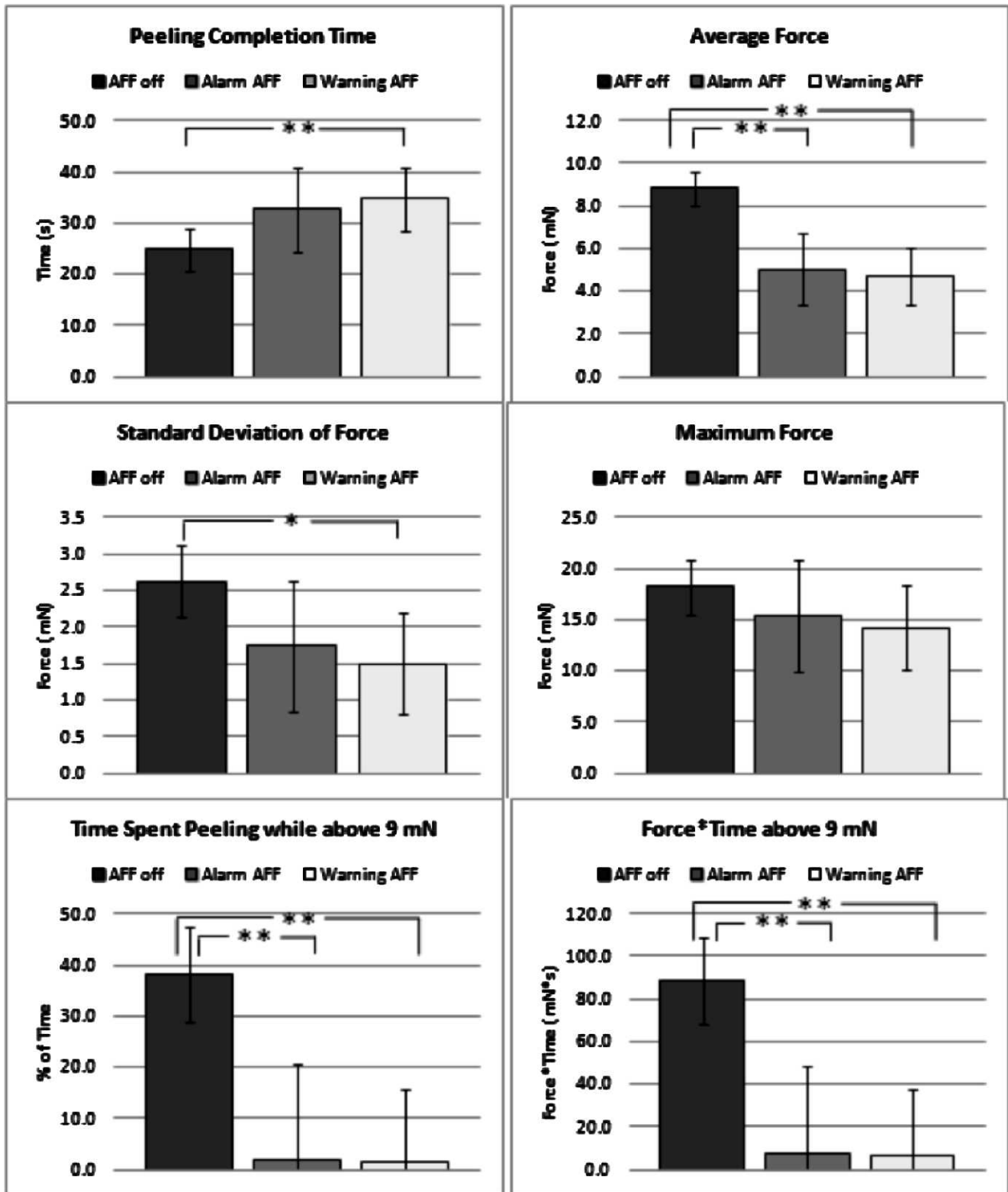


FIGURE 5. Comparisons of peeling completion time, average force, SD of force, maximum force, percentage of time spent peeling while above 9 mN, and force \times time above 9 mN between peels for all participants who performed the first five membrane peels without AFF, with alarm AFF, and with warning AFF. Capped lines indicate 95% confidence intervals from regression models. * $P < 0.05$, ** $P < 0.01$.

TABLE 1. Effect of AFF on Epiretinal Membrane Peeling Performance

Variable	AFF Off #1, <i>n</i> = 80	AFF Off #2, <i>n</i> = 80	<i>P</i> Value
Peeling completion time, s	24.8 (21.2, 28.4)	39.0 (35.4, 42.6)	<0.001
Average force, mN	8.8 (8.1, 9.5)	5.2 (4.5, 5.9)	<0.001
SD of force, mN	2.6 (2.2, 3.0)	1.1 (0.7, 1.5)	<0.001
Maximum force, mN	18.2 (16.1, 20.3)	10.0 (7.9, 12.2)	<0.001
Percentage of time spent peeling above 9 mN	38.2 (30.9, 45.5)	2.3 (0.0, 9.6)	<0.001
Force × time above 9 mN, mN·s	88.7 (71.1, 106.2)	8.1 (0.0, 25.7)	<0.001

For peels with AFF turned off, comparisons of peeling completion time, average force, SD of force, maximum force, percentage of peeling spent peeling while above 9 mN, and force × time above 9 mN between the first five peels performed at the beginning of the session (AFF off #1) and the last five peels at the end of the session (AFF off #2). 95% confidence intervals from regression models are shown in parentheses.

× 776 Color Flea2, C-mount, Sony ICX204 CCD; Sony Corp., Tokyo, Japan).

Epiretinal Membrane Phantom

We simulated epiretinal membranes with a phantom membrane composed of the sticky tabs on 19-mm latex-free, waterproof bandages (Clear Bandages; Rite Aid Corp., Camp Hill, PA) as described in Balicki et al.¹¹ The tabs were cut into 3-mm-wide × 20-mm-long strips and adhered to a stable platform with double-sided tape. The sticky bandage tabs were then peeled from their backing, which remained affixed to the platform by the double-sided tape.

Ten new phantom membranes were placed for each participant. Since the phantoms were often more adherent during their first peel, each phantom was prepped for a new participant by being peeled ten times. Between each peeling, the phantoms were lightly brushed back onto the platform three times with a nylon fiber brush. This process allowed for consistency from the phantoms between peels. The platform was marked with start and end lines spaced 13 mm apart, with the beginning edge free from the platform so that the participants could easily grab the phantom membrane (Fig. 2).

Auditory Force Feedback

Real-time AFF was provided to the participants to inform them of how much force was being applied while peeling the phantom membranes. The audio feedback consisted of “beeps” played at varying frequencies, depending on the level of force being sensed by the tool. We chose levels based on research done in porcine cadaver eyes, which showed most forces during vitreoretinal surgery are at or below 7.5 mN.^{2,14} Two modes of AFF were used, “alarm” and “warning.” In alarm AFF mode, the system was silent until a force of 9 mN was reached. At forces of 9 mN and higher, a continuous high-pitch tone sounded to indicate excessive and potentially damaging forces were being applied to the tissue. In warning AFF mode, the system emitted graded sounds based on the level of force. For forces < 3 mN, the audio was silent. At 3 mN, the system emitted beeps at a frequency of 1 Hz that proportionately increased in frequency as the force rose from 3 to 9 mN. The graded warning beeps alerted the participant that he/she was approaching the upper limit of safe forces. At a force of 9 mN and higher, the graded beeps stopped and the system emitted the same continuous high-pitch tone used in the alarm AFF mode.

Membrane Peeling Trials Protocol

The surgical task consisted of peeling the phantom membrane off of the platform with the force-sensing microforceps as quickly as possible without exceeding 9 mN of force. Before the first peel, each participant watched a demonstration video of a peel performed without exceeding 9 mN of force. Before each peel, the force sensor was zeroed while the participant closed the forceps. The membrane peeling was visualized

in stereo through an operating microscope (Zeiss OPMI CS-NS surgical microscope; Carl Zeiss AG, Oberkochen, Germany) with an S4 ceiling attachment. The zoom was placed at 0.5 for all participants. Participants began peeling when they heard a tone and stopped peeling when they reached the end target line (a distance of 13 mm) or after 45 seconds had elapsed. The participants were instructed to keep the instrument 90° relative to the direction of peeling, to minimize the unmeasured force along the tool axis.

The study was done in two parts: arms A and B. Arm A consisted of 10 sequential peels performed with alarm AFF mode and 10 sequential peels with warning AFF mode. Participants ID 1 through 11 were randomly selected within experience levels to peel first with the alarm AFF or the warning AFF (Fig. 3). Arm B was designed to record forces during peels with no AFF. For participants ID 12 through 28, each participant first performed five peels with no AFF (AFF off #1), then 10 sequential peels with warning AFF and 10 sequential peels with alarm AFF, the order randomized within experience levels, and finished with five peels with no AFF (AFF off #2), for a total of 30 peels. For each peel, the recorded parameters included measured instrument forces, instrument tip position, distance peeled, and task completion time. Data from participant ID number 22 (an attending surgeon) could not be included in the analysis due to a computer error that prevented any of the data from being recorded.

Task Performance Measures

Plotting the peeling trials with force against time illustrated the differences in peeling performance based on the absence or presence of AFF (Fig. 4). After analyzing the force plots, we chose six quantifiable parameters, to assess the effects of AFF on task performance. Those parameters included: peeling completion time, maximum force, average force, SD of force, percentage of time spent peeling above 9 mN, and force × time above 9 mN. Peeling completion time was included since prior studies have shown mixed results as to the effect of adding haptics and force feedback substitution on procedural times.^{4,8,15} The average force applied while peeling and the SD of that average force were used to show the control and precision, respectively, of the participant. To demonstrate the effect of AFF on reducing the most damaging forces, the maximum force for each peel was included in the analysis. Since each participant was tasked to keep the force at the tool tip below 9 mN, we also directly measured the percentage of time spent peeling above 9 mN. Finally, to measure total amount of unwanted force being applied during each peeling, we calculated the force × time area above 9 mN of each force plot.

Survey

A written qualitative subjective survey was completed by each participant at the conclusion of the peeling trials to assess whether alarm AFF or warning AFF was preferred and the reason(s) for that preference.

TABLE 2. Peeling Performance with AFF by Surgeon Experience

Variable	Medical Students			Residents			Fellows and Attendings		
	AFF Off #1, <i>n</i> = 10	AFF Off #2, <i>n</i> = 10	<i>P</i> Value	AFF Off #1, <i>n</i> = 30	AFF Off #2, <i>n</i> = 30	<i>P</i> Value	AFF Off #1, <i>n</i> = 40	AFF Off #2, <i>n</i> = 40	<i>P</i> Value
Peeling completion time, s	24.9 (14.9, 35.0)	40.5 (30.4, 50.5)	<0.001	24.1 (18.3, 29.9)	41.3 (35.5, 47.1)	<0.001	25.3 (20.2, 30.3)	36.9 (31.9, 41.9)	<0.001
Average force, mN	8.9 (7.0, 10.8)	4.3 (2.4, 6.2)	<0.001	9.7 (8.6, 10.8)	5.4 (4.3, 6.5)	<0.001	8.2 (7.2, 9.1)	5.2 (4.3, 6.2)	<0.001
SD of force, mN	2.5 (1.5, 3.6)	0.9 (0.0, 1.9)	0.012	3.3 (2.7, 3.9)	1.2 (0.6, 1.8)	<0.001	2.1 (1.6, 2.7)	1.1 (0.6, 1.7)	0.002
Maximum force, mN	18.5 (13.2, 23.9)	7.9 (2.5, 13.3)	<0.001	21.7 (18.6, 24.8)	11.4 (8.3, 14.5)	<0.001	15.5 (12.8, 18.2)	9.6 (6.9, 12.3)	<0.001
Percentage of time spent peeling above 9 mN	35.1 (14.5, 55.7)	0.0 (0.0, 20.6)	<0.001	42.3 (30.4, 54.2)	1.7 (0.0, 13.6)	<0.001	35.9 (25.6, 46.2)	3.4 (0.0, 13.7)	<0.001
Force × time above 9 mN, mN·s	72.6 (25.3, 120.0)	0.0 (0.0, 47.4)	0.007	115.3 (88.0, 142.6)	7.8 (0.0, 35.1)	<0.001	72.7 (49.1, 96.4)	10.4 (0.0, 34.1)	<0.001

For peels with AFF turned off, comparisons of peeling completion time, average force, SD of force, maximum force, percentage of peeling spent peeling while above 9 mN, and force × time above 9 mN between the first five peels performed at the beginning of the session (AFF off #1) and the last five peels at the end of the session (AFF off #2). Stratified by participant's surgical skill level. 95% confidence intervals from regression models are shown in parentheses.

Statistical Analysis

The regression analyses were conducted using generalized, linear mixed effects models, with a random participant intercept to account for correlation among the performance outcome within the same participant. Statistical comparisons among no AFF, alarm AFF, and warning AFF were evaluated with linear combinations of the estimates based on the mixed models. Values of $P \leq 0.05$ were considered statistically significant. All analyses were performed using commercial statistical software (Stata 11.2; StataCorp, College Station, TX). Analysis of responses to the qualitative survey was conducted using a single-factor ANOVA (Microsoft Excel; Microsoft Corp., Redmond, WA).

RESULTS

AFF Improves Performance of Unfamiliar Tasks

The first goal was to assess whether AFF improves performance outcomes for participants performing unfamiliar peeling tasks. To assess this effect, the first five peels performed by participants were analyzed ($n = 80$ for AFF off, $n = 20$ for alarm AFF, and $n = 35$ for warning AFF). There was a significant difference between no AFF and warning AFF for all performance parameters except maximum force (Fig. 5). Differences between no AFF and alarm AFF were significantly different in three of the performance parameters including: average force, percentage of time spent peeling above 9 mN, and force × time above 9 mN. Compared with alarm AFF, peels performed with warning AFF showed nonsignificant trends in the first five peels of longer completion time and lower values for the remaining five performance parameters.

Training with AFF Improves Performance

Next, to determine whether training with AFF improves peeling performance once AFF is turned off, a comparison was made for participants in study arm B between the five peels performed without AFF at the beginning and the last five peels performed without AFF at the end of each participant's peeling trials. AFF training resulted in a significant increase in peeling completion time and a significant reduction in the other five performance parameters (Table 1). These significant differences remained for all measured variables when stratified by experience level (Table 2).

AFF Continues to Improve Performance, Even after Training

After demonstrating that a short training session with AFF can improve peeling performance even once AFF is turned off, we wanted to determine whether the now trained participant would still benefit from AFF. To determine whether AFF improves peeling performance beyond training alone, we compared alarm AFF, warning AFF, and the last five peels performed without AFF among study arm B participants ($n = 80$ for AFF off #2, $n = 160$ for alarm, and $n = 160$ for warning). Compared with AFF off, warning AFF had significantly longer completion time and significantly less average force, SD of force, percentage of time spent peeling above 9 mN, and force × time above 9 mN (Fig. 6). Compared with AFF off, alarm AFF also had significantly longer completion time and significantly lower average force. Compared with alarm AFF, peels performed with warning AFF had significantly lower SD of force, maximum force, percentage of time spent peeling above 9 mN, and force × time above 9 mN.

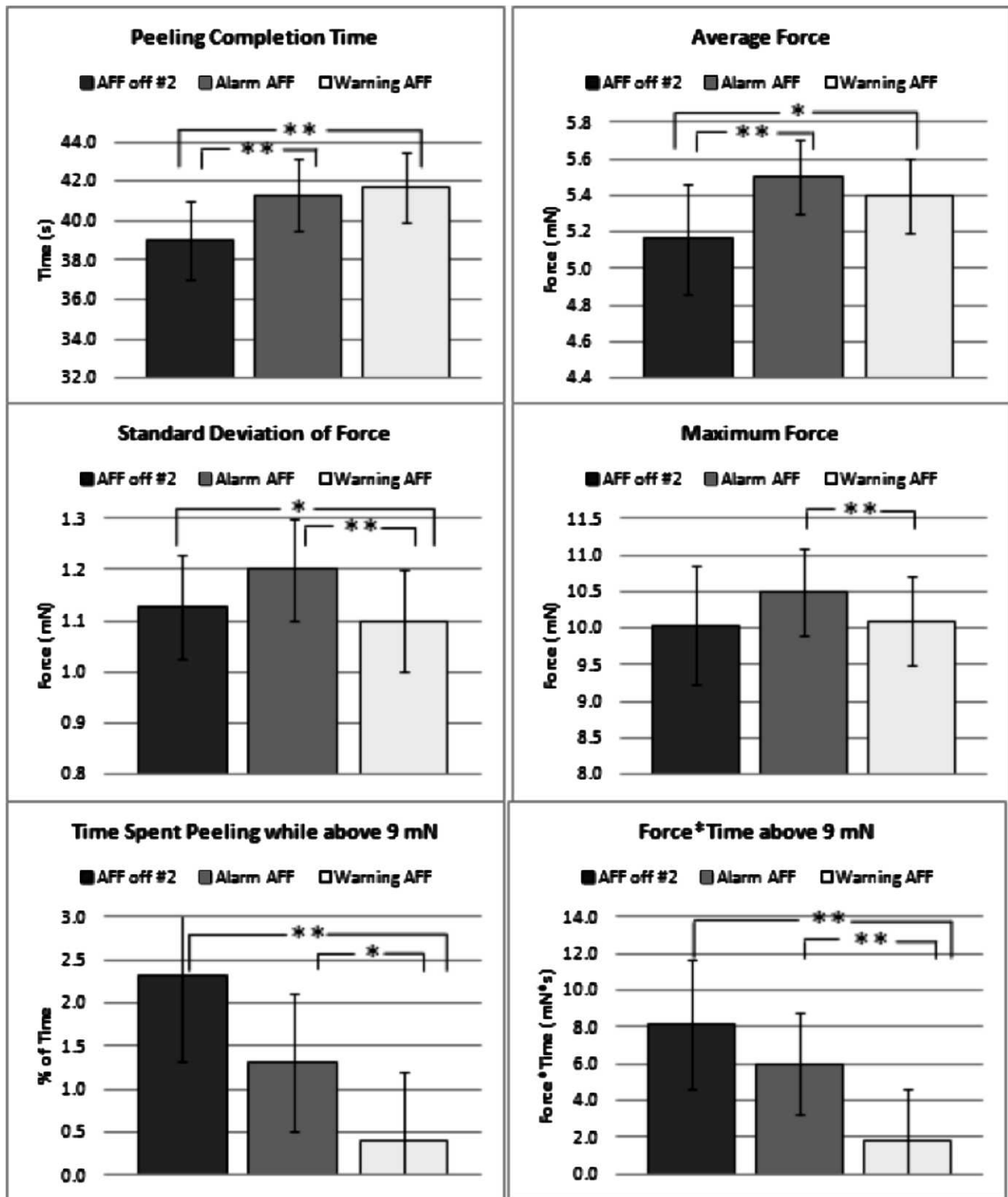


FIGURE 6. Comparisons of peeling completion time, average force, SD of force, maximum force, percentage of time spent peeling while above 9 mN, and force \times time above 9 mN between peels for all participants who performed membrane peeling without AFF during the last five peels of the trial (AFF off #2), with alarm AFF, and with warning AFF. Capped lines indicate 95% confidence intervals from regression models. * $P < 0.05$, ** $P < 0.01$.

TABLE 3. Comparison of Warning AFF and Alarm AFF on Peeling Performance

Variable	Alarm AFF, <i>n</i> = 270	Warning AFF, <i>n</i> = 270	Alarm vs. Warning <i>P</i> Value
SD of force, mN	1.3 (1.2, 1.5)	1.2 (1.0, 1.3)	0.048
Maximum force, mN	11.6 (10.5, 12.8)	10.8 (9.6, 11.9)	0.045
Peeling completion time, s	39.3 (37.4, 41.3)	40.0 (38.0, 41.9)	0.254
Average force, mN	5.3 (5.1, 5.6)	5.2 (4.9, 5.4)	0.189
Percentage of time spent peeling while above 9 mN	1.6 (0.0, 3.6)	0.7 (0.0, 2.7)	0.482
Force × time above 9 mN, mN-s	7.0 (1.8, 12.2)	3.1 (0.0, 8.3)	0.221

For all participants, comparisons of SD of force, maximum force, peeling completion time, average force, percentage of time spent peeling while above 9 mN, and force × time above 9 mN between peels performed with alarm AFF and with warning AFF. 95% confidence intervals from regression models are shown in parentheses.

Warning AFF Improves Performance More than Alarm AFF

Finally, to determine whether one type of auditory feedback is better than another, we compared warning AFF to alarm AFF for every participant. Peels performed with warning AFF had significantly lower SD of force and maximum force than peels performed with alarm AFF (Table 3). Compared with alarm AFF, warning AFF additionally showed nonsignificant trends for lower average force, percentage of time spent peeling above 9 mN, and force × time above 9 mN. When stratified by training level, there were no significant differences between peels performed with alarm AFF and warning AFF (Table 4).

Qualitative Survey Results

AFF with warning was preferred overall by 75% of the participants since it allowed more precise peeling at a consistent pace and greater opportunity to react to rising forces (Fig. 7). For those that preferred alarm AFF, the additional beeps present in the warning mode were found to be a source of distraction. Many commented that having AFF available for residency training would be valuable. Differences between training levels were insignificant ($P = 0.26$, ANOVA).

DISCUSSION

Although many studies have explored the addition of haptic feedback and force feedback substitution to laparoscopic and

robotic surgery, few have looked at force feedback substitution in microsurgical procedures.^{3,4,7-9,15} The goal of adding force feedback to minimally invasive surgery is to regain tactile sensation lost through the detachment created by laparoscopic instruments. The overall effects of adding force feedback to such cases are ambiguous, but greater effect has been shown when adding haptic force feedback, visual force feedback, and AFF substitution to robotic surgery, a form of surgery where there is currently not any form of force feedback.⁴ In contrast to minimally invasive surgery, the addition of force feedback substitution to ophthalmic procedures is not one of regaining lost feedback, but the addition of force feedback to a system that never had it. This study builds on work previously done by our group toward the development of “smart” microsurgical instruments designed to improve surgical performance through detection of forces that are below human tactile sensation.^{1,10,11,16,17} To the best of our knowledge, this is the first prospective trial exploring the effects of adding AFF to simulated ophthalmic surgery.

In this study, we demonstrated that AFF can effectively improve peeling performance when participants are asked to perform an unfamiliar procedure. The only training that our participants received prior to performing their first peels was to view a video of a peel done with a high level of precision at a consistent peeling force < 9 mN. This is representative of training in ophthalmic surgery where the surgeon must often rely on visual cues such as tissue deformation instead of haptic feedback. Variances between individual eyes result in a new peeling experience for every procedure, illustrating the importance of being able to adapt to new tasks and the

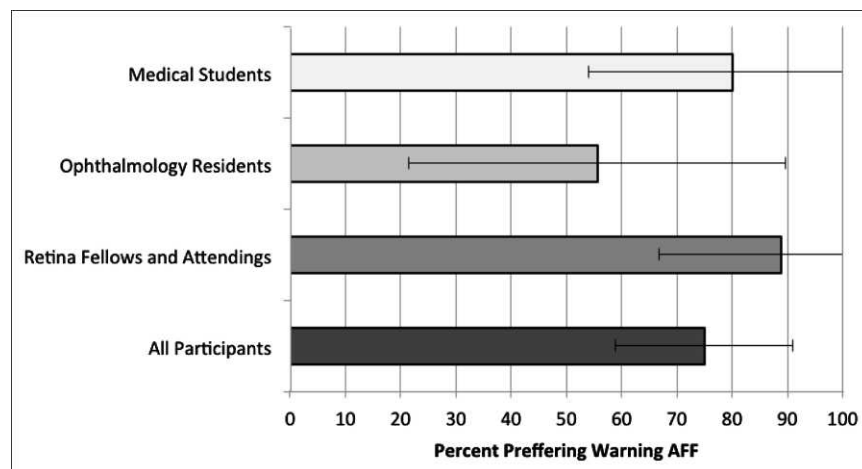


FIGURE 7. Qualitative survey results showing percentage of participants who preferred the warning AFF over alarm AFF. Capped lines indicate 95% confidence intervals.

TABLE 4. Comparison of Warning AFF and Alarm AFF on Peeling Performance Stratified by Surgeon Experience

Variable	Medical Students			Residents			Fellows and Attendings		
	Alarm AFF, n = 100	Warning AFF, n = 100	Alarm vs. Warning P Value	Alarm AFF, n = 90	Warning AFF, n = 90	Alarm vs. Warning P Value	Alarm AFF, n = 80	Warning AFF, n = 80	Alarm vs. Warning P Value
SD of force, mN	1.5 (1.3, 1.8)	1.4 (1.1, 1.6)	0.21	1.2 (1.0, 1.5)	1.1 (0.9, 1.4)	0.49	1.2 (0.9, 1.4)	0.9 (0.7, 1.2)	0.12
Maximum force, mN	13.8 (12.2, 15.3)	12.8 (11.3, 14.3)	0.16	10.6 (9.0, 12.2)	10.3 (8.7, 11.9)	0.65	10.1 (8.4, 11.8)	8.8 (7.1, 10.5)	0.10
Peeling completion time, s	37.9 (34.7, 41.0)	39.2 (36.0, 42.3)	0.18	39.0 (35.7, 42.3)	40.0 (36.6, 43.3)	0.33	41.5 (37.9, 45.0)	41.0 (37.5, 44.5)	0.68
Average force, mN	5.2 (4.8, 5.6)	5.0 (4.6, 5.4)	0.48	5.5 (5.0, 5.9)	5.2 (4.8, 5.6)	0.24	5.3 (4.9, 5.8)	5.2 (4.8, 5.7)	0.69
Percentage of time spent peeling while above 9 mN	2.4 (0.0, 5.7)	1.2 (0.0, 4.5)	0.54	1.3 (0.0, 4.8)	0.6 (0.0, 4.1)	0.75	0.9 (0.0, 4.6)	0.3 (0.0, 4.0)	0.79
Force × time above 9 mN, mN·s	10.5 (2.2, 18.7)	5.0 (0.0, 13.3)	0.30	5.6 (0.0, 14.3)	2.6 (0.0, 11.3)	0.59	4.4 (0.0, 13.6)	1.2 (0.0, 10.5)	0.60

For all participants, comparisons of SD of force, maximum force, peeling completion time, average force, percentage of time spent peeling while above 9 mN, and force × time above 9 mN between peels performed with alarm AFF and with warning AFF. Stratified by participant's surgical skill level, 95% confidence intervals from regression models are shown in parentheses.

potential benefit of force-sensing tools. Additionally, as surgical training programs continue to adapt curricula to the new Accreditation Council for Graduate Medical Education work hour restrictions, training with force-sensing tools could help new residents acquire microsurgical skills more quickly.

AFF did more than just improve peeling performance at the beginning of each participant's set of peeling trials. By comparing the peels done without AFF at the end of each participant's set to those done with warning AFF, we showed that even after training with AFF for 20 peels, performance continues to be better with AFF. Additionally, it was shown that even short training sessions with AFF can improve subsequent peeling performance when the AFF is turned off. As a result, even if participants would choose not to have the extra auditory sounds during actual surgery, training with the force feedback system in a "wet" lab could still have a positive impact on surgical outcomes.

We were surprised to find in our study that, for the most part, differences between peeling with and without AFF were not influenced by experience level. We had suspected that AFF would have shown more improvements in the less experienced participants, as was seen by Reiley et al.,⁸ where experienced da Vinci surgeons showed no difference in performance when provided with active force feedback substitution. Although the more experienced participants in our study did perform the tasks better than those with less surgical experience, they still achieved greater precision and control with AFF. We acknowledge, however, that the lack of influence by experience level could reflect a limitation of our epiretinal membrane phantom because the peeling movement was not restricted by a pivot point, as required when operating through a sclerotomy.

A concern that we have had during the development of feedback substitution is how intrusive such a system would be to the surgeon and the operating environment as a whole. Because we were unable to decide whether a continuous feedback system such as the warning AFF or a simple alarm system would be most effective, we included both in this study. Although the warning AFF was significantly better than the alarm AFF in SD of force and maximum force, it was not significantly different from alarm AFF in the other four measured performance parameters. However, the insignificant trends for warning AFF to be better than alarm AFF became significant when comparing AFF to peels performed without AFF. Warning AFF outperformed no AFF in nearly all comparisons, whereas alarm AFF only sometimes achieved significantly better performance than no AFF. Additionally, despite the insignificant head-to-head differences between alarm AFF and warning AFF, our qualitative subjective survey showed that most participants preferred the warning AFF, independent of training level. Many participants thought the graded sounds of the warning AFF conferred confidence and appreciated the additional real-time feedback. Many of the ophthalmology residents participating in the trial commented that they wished such a device was currently available for their training.

Given the technical challenges associated with testing a new surgical instrument, our study had several limitations. We had considered testing the instrument on more advanced phantoms or even cadaveric eyes, but this would have limited the number of peels per participant and would have introduced additional uncontrolled variables. Moreover, even though the phantom membrane, the influences on surgical movement imposed by a sclerotomy, and the procedural task of peeling in a straight line were quite different from the ophthalmic procedures they were meant to model, the psychomotor coordination required to complete the task is

representative of the skills required to perform ophthalmic surgery. A major limitation of the current instrument is its ability to only measure forces in the x - and y -planes, but not the z -plane (tool axis). Axial forces were limited in this study by using a simple, single-dimension peeling task. Since this study was completed, a prototype force-sensing tool has been developed by our group with a full 3DOF.

The number of participants, although similar to many prior studies comparing force feedback substitution, was relatively small. Additionally, to account for the effect of the different order of modes interfering with each participant's learning, we initially compared only the first five peels for each participant, which further reduced the sample size. However, a strength of this study was the large number of peels performed by each participant, giving us the best estimate of each participant's true performance. A potential drawback of AFF substitution is the potential increase in procedure completion time. However, in a series of complex tasks such as vitreoretinal surgery, it is possible that the improved precision of each individual movement will ultimately neutralize any potential increase in performance time. It will be important to see in more realistic models whether such increases in time will have clinically significant effects. A final limitation was that each participant performed all the peels consecutively during one session. To truly assess the benefit of training with AFF, participants would need to be reevaluated using the same peeling task after an interval period of time.

We are planning future studies to incorporate 3DOF force-sensing microsurgical instruments under more realistic phantom tissues to more accurately simulate surgical task performance. Additionally, recent work by our group has shown that it might not be the absolute level of force that is most important when trying to control tissue damage, but the rate of force generation.¹ We plan to continue to characterize the forces required to damage tissue during ophthalmic procedures and whether limiting that force can improve surgical outcomes. Our group has also designed a sterilizable force-sensing microforceps that was manufactured using a modular design.¹⁸ FGB strain-sensing fibers, which are sterilizable, were glued to a nitinol tube that is attached to a disposable microforceps tip at the distal end and a disposable handle apparatus at the proximal end. Although the microforceps components are disposable, the nitinol tube with the FGB fibers can be detached and then sterilized after use.

Our results suggest that the addition of AFF substitution reduces excessive force application and improves surgical precision during repetitive ophthalmic phantom membrane peeling, regardless of surgical experience. Thus, new force-sensing surgical tools with AFF substitution such as the one presented here, have potential applications for surgical training and future use in vivo to improve surgical outcomes.

References

- Sunshine S, Balicki M, He X, et al. A force-sensing microsurgical instrument that detects forces below human tactile sensation. *Retina*. 2013;33:200–206.
- Gupta PK, Jensen PS, de Juan E. Surgical forces and tactile perception during retinal microsurgery. In: *Medical Image Computing and Computer-Assisted Intervention MICCAI'99, Second International Conference*. Vol. 4. Morrisville, NC: MICCAI Society; 1999:1218–1225.
- Kontarinis DA, Howe RD. Tactile display of vibratory information in teleoperation and virtual environments. *Presence*. 1995;4:387–402.
- van der Meijden O, Schijven M. The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review. *Surg Endosc*. 2009;23:1180–1190.
- Bethea BT, Okamura AM, Kitagawa M, et al. Application of haptic feedback to robotic surgery. *J Laparoendosc Adv Surg Tech A*. 2004;14:191–195.
- Okamura AM, Smaby N, Cutkosky MR. An overview of dexterous manipulation. In: *Proceedings of the IEEE International Conference on Robotics and Automation, 2000. ICRA '00*. Vol. 1. Piscataway, NJ: IEEE; 2000:255–262.
- Horeman T, Rodrigues SP, van den Dobbelen JJ, Jansen FW, Dankelman J. Visual force feedback in laparoscopic training. *Surg Endosc*. 2012;26:242–248.
- Reiley CE, Akinbiyi T, Burschka D, Chang DC, Okamura AM, Yuh DD. Effects of visual force feedback on robot-assisted surgical task performance. *J Thorac Cardiovasc Surg*. 2008;135:196–202.
- Panait L, Akkary E, Bell RL, Roberts KE, Dudrick SJ, Duffy AJ. The role of haptic feedback in laparoscopic simulation training. *J Surg Res*. 2009;156:312–316.
- Iordachita I, Sun Z, Balicki M, et al. A sub-millimetric, 0.25 mN resolution fully integrated fiber-optic force-sensing tool for retinal microsurgery. *Int J Comput Assist Radiol Surg*. 2009;4:383–390.
- Balicki M, Uneri A, Iordachita I, Handa J, Gehlbach P, Taylor R. Micro-force sensing in robot assisted membrane peeling for vitreoretinal surgery. *Med Image Comput Comput Assist Interv*. 2010;13:303–310.
- He X, Balicki MA, Kang JU, et al. Force sensing micro-forceps with integrated fiber Bragg grating for vitreoretinal surgery. In: *Optical Fibers and Sensors for Medical Diagnostics and Treatment Applications XII*. Vol. 8218. *Proc SPIE* 2012: 82180W.
- Kazanides P, DiMaio S, Deguet A, et al. The Surgical Assistant Workstation (SAW) in minimally-invasive surgery and microsurgery. In: *MICCAI Workshop on Systems and Architecture for Computer Assisted Interventions*. Morrisville, NC: MICCAI Society; 2010.
- Jagtap AS, Riviere CN. Applied force during vitreoretinal microsurgery with handheld instruments. *Conf Proc IEEE Eng Med Biol Soc*. 2004;4:2771–2773.
- MacFarlane M, Rosen J, Hannaford B, Pellegrini C, Sinanan M. Force-feedback grasper helps restore sense of touch in minimally invasive surgery. *J Gastrointest Surg*. 1999;3:278–285.
- Liu X, Iordachita II, He X, Taylor RH, Kang JU. Miniature fiber-optic force sensor based on low-coherence Fabry-Pérot interferometry for vitreoretinal microsurgery. *Biomed Opt Express*. 2012;3:1062–1076.
- Uneri A, Balicki MA, Handa J, Gehlbach P, Taylor RH, Iordachita I. New steady-hand eye robot with micro-force sensing for vitreoretinal surgery. In: *Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatron*. 2010;26–29:814–819.
- Kuru I, Gonenc B, Balicki M, et al. Force sensing micro-forceps for robot assisted retinal surgery. In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2012:1401–1404.