Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Virtual Reality Warm-up Before Robot-assisted Surgery: A Randomized Controlled Trial

Jason D. Kelly, PhD,[a,*] *Timothy M. Kowalewski, PhD,*[a] *Tim Brand, MD,*[b] *Anna French, PhD,*[a] *Michael Nash, MS,*[b] *Lois Meryman, BS,*[b] *Nicholas Heller, BS,*[a] *Nancy Organ, MS,*[b] *Evalyn George, BS,*[c] *Roger Smith, MD,*[d] *Mathew D. Sorensen, MD,*[b] *Bryan Comstock, MS,*[b] *and Thomas S. Lendvay, MD, FACS*[b]

[a] *University of Minnesota, Department of Mechanical Engineering, Minneapolis, Minnesota*
[b] *University of Washington, Department of Urology, Seattle, Washington*
[c] *Madigan Army Medical Center, Henry M. Jackson Foundation, Tacoma, Washington*
[d] *Florida Hospital Nicholson Center, Orlando, Florida*

ARTICLE INFO

Article history:
Received 10 August 2020
Received in revised form
20 January 2021
Accepted 22 January 2021
Available online 31 March 2021

Keywords:
Virtual reality
Warm-up robotic surgery readiness

ABSTRACT

Trial design: This was a randomized controlled trial.
Background: Intraoperative errors correlate with surgeon skill and skill declines with intervals of inactivity. The goals of this research were to identify the optimal virtual reality (VR) warm-up curriculum to prime a surgeon's technical skill and validate benefit in the operating room.
Materials and methods: Surgeons were randomized to receive six trial sessions of a designated set of VR modules on the da Vinci Skills Simulator to identify optimal VR warm-up curricula to prime technical skill. After performing their curricula, warm-up effect was assessed based on performance on a criterion task. The optimal warm-up curriculum was chosen from the group with the best task time and video review–based technical skill. Robot-assisted surgery–experienced surgeons were then recruited to either receive or not receive warm-up before surgery. Skill in the first 15 min of surgery was assessed by blinded surgeon and crowdworker review as well as tool motion metrics. The intervention was performing VR warm-up before human robot-assisted surgery. Warm-up effect was measured using objective performance metrics and video review using the Global Evaluative Assessment of Robotic Skills tool. Linear mixed effects models with a random intercept for each surgeon and nonparametric modified Friedman tests were used for analysis.
Results: The group performing only a Running Suture task on the simulator was on average 31.3 s faster than groups performing other simulation tasks and had the highest Global Evaluative Assessment of Robotic Skills scores from 41 surgeons who participated. This was chosen as the optimal curriculum. Thereafter, 34 surgeons completed 347 surgeries with corresponding video and tool motion data. No statistically significant differences in skill were observed with the warm-up intervention.
Conclusions: We conclude that a robotic VR warm-up before performing the early stages of surgery does not impact the technical skill of the surgeon.
© 2021 Elsevier Inc. All rights reserved.

* *Corresponding author*. 16 Hull Street Apt. 3, Boston, MA, 02113. Tel.: +1 731-446-1324; fax: +1 731-446-1324.
E-mail address: kell1917@umn.edu (J.D. Kelly).
0022-4804/$ – see front matter © 2021 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.jss.2021.01.037

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

Introduction

Medical errors account for 98,000 deaths each year, atop a higher number of patient complications.[1] An estimated one-third of these errors are surgical, and the average American can expect to have seven surgeries in their lifetime.[2-4] Although surgical simulation's role in off-line teaching is accepted,[5] simulation may also have a role in identifying decrements in surgical performance and offer a system for mitigating skill decay.[6-9] Perez *et al*. concluded that skill may decay at different degrees depending on the type of skill and that training to mitigate decay may require targeted curricula based on the specific decay signatures.[7] These methods were described for laparoscopic surgery skills, yet no warm-up has been studied for robot-assisted surgery which has seen considerable adoption in the United States with the creation of the da Vinci surgical robot by Intuitive Surgical (Sunnyvale, CA).[8]

Although several other checklist-type preparation procedures exist in the operating room, there is still no metric which measures how prepared a surgeon is after periods of inactivity.[9] In other professions, after participating in a warm-up, task performance can be elevated.[10] Previous work additionally shows that a warm-up decrement exists, in which the effects of warm-up begin decaying approximately 30 min after performing a warm-up task.[10-13] These hypotheses have already found four factors which most affect warm-up decrement: (1) intervals between practice, (2) complexity and type of task, (3) strategies used to maintain skills, and (4) individual differences.[14,15] Virtual reality (VR) proficiency-based warm-up modules have been shown to most effectively combat skill decay.[16-20] This research group identified in a prior study that VR warm-up primes experienced robotic surgeons when looking at task time (TT) and errors; thus, we sought to assess the warm-up benefit among practicing surgeons.[19]

The da Vinci Skills Simulator (dVSS) replaces the physical robotic arms of a da Vinci robot with virtual representations of the tools to simulate a surgical environment (Fig. 1).[21] The simulator tasks exercise a surgeon's ability to move objects, target camera position, dissect tissues, and suture.

Measuring surgical technical skill in an objective manner is difficult, with the gold standard being video review by expert surgeons. Previous work has shown that video review of technical skill by nonexpert crowds is accurate. When compared with expert surgeon review, crowds of nonexperts were able to predict the pass and fail rate of surgeons with 100% accuracy in retrospective studies.[22,23] Tool motion data captured from the da Vinci have also been associated with surgical skill. Metrics, such as TT, spectral arc length (SAL), camera use per minute (CUP), normalized angular displacement (NAD), rate of orientation change (ROC), and mean velocity (MV), have all been found to associate with surgical skill.[24-28] Of these objective metrics, TT has repeatedly been shown to be the most accurate measure in classifying technical surgical skill, when compared with the gold standard.[23,27]

In this study, we sought to find the optimal VR warm-up curriculum to properly prime surgeons for robot-assisted surgery and validate the warm-up benefit in the operating room by measuring objective performance metrics and obtaining video review of participants.

Fig. 1 – The da Vinci Skills Simulator allows surgeons to practice typical surgical tasks outside of real surgery. (Color version of the figure is available online.)

Materials and methods

Aim 1

Participant recruitment

Surgeons and trainees were recruited after giving informed consent to an institutional review board–approved (#41730) study. Participating surgeons were from urology, gynecology, and general surgery at four clinical sites including military, Veteran's Administration Hospital, public academic, and private medical centers. The trial was a crossover design with a 1 to 1 allocation ratio. Each of the participants was asked to complete a standard demographic questionnaire to learn their experience and any factors impacting responses to warm-up simulations. The information obtained included typical demographics info such as age, gender, and so forth as well as handedness, experience in surgery/robotic surgery, musical experience, and experience playing video games. As the participants in this study were surgeons, this was not a registered clinical trial.

dVSS curricula

Modules used for surgeon priming needed to be technically challenging to stimulate the surgeon in a way that most closely corresponds to movements present in a surgical environment. The criteria used to select modules were as follows: (1) prior validation in the literature, and/or 2) involvement of multiple skills simultaneously, and/or 3) containing content that closely simulates actual surgery. Following these criteria, we chose to use the "Ring and Rail 2"

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

and "Match Board 3" modules from the "EndoWrist Manipulation 2" category (Fig. 2). "Ring and Rail 2" has previously been found to effectively discriminate surgical skills, and "Match Board 3" has been found to be the most difficult of the dVSS modules among experienced surgeons.[17,29] From the "Needle Driving" category, "Suture Sponge 3" was chosen as it has been found to vary greatly between experienced, intermediate, and novice performers.[29] "Running Suture 2" was the final module chosen and also used as the criterion module against which all groups of participants were tested for proficiency, as it most closely represents the actions in a real surgical suturing procedure. Each of these four modules can be completed in approximately 2-3 min, making them short enough to be feasibly used as a component of a final curriculum.

Proficiency testing

To normalize the skills of participants, each underwent proficiency training before being randomized to one of five curriculum groups (Fig. 3) consisting of modules shown in literature to improve technical skill. Proficiency testing included completion of the Intuitive Surgical da Vinci didactic web-based curriculum. To obtain benchmark performance goals for the proficiency training, two expert robot-assisted surgeons completed the four modules used until they made no errors. The average time from the best TT for each surgeon was used as the benchmark. Participants were deemed 'proficient' once they met TT benchmarks for three consecutive trials.

The four dVSS modules were split into five curriculum groups with different combinations of modules to test which curriculum was most efficient for priming surgeons. The modules in each of these five groups are shown in Table 1. Within each group, half were randomized to perform the tasks monthly and half performed biweekly. After 8 wk of performing at one frequency, the participants were then crossed over to the opposing frequency group for a total of six trial sessions, to allow for certain baselines in statistical models to control for intervals of inactivity. After performing his/her respective curriculum, each participant performed the criterion module (Running Suture 2).

Video review

Video was obtained from each criterion VR task session, and two blinded expert surgeons reviewed videos of the performances and rated them using the Global Evaluative Assessment of Robotic Skills (GEARS) tool.[30] The 'Autonomy' domain of the tool was removed as it could not be used with a post-performance video review, leaving a total possible score of 5-25.

Statistical models

The primary goals of aim 1 were to evaluate (1) the magnitude of skill decay between the two standardized assessment intervals (monthly and biweekly) and (2) the curriculum which optimizes the participant's GEARS scores and results in the smallest deviation from performer baseline performance.

We used a linear mixed effects model which accounts for interval assessment period, timing of assessment, and contrast between the different curricula. This model included a random intercept to account for clustering of study participants and a fixed effect to adjust for assessment time between different trials. To account for the average percentage of skill decay between groups, the contrast in GEARS scores between groups performing monthly and groups performing biweekly was accounted for in the model. The model did not take into account differences in participant demographics when computing a result. The module which yielded the highest average combined GEARS score after accounting for all of these factors was used as the optimal warm-up curriculum for aim 2.

Aim 2

Participant recruitment

Aim 1 practicing surgeons were allowed to participate in aim 2 of the study in addition to newly recruited surgeons. All newly recruited surgeons consented to participating in this institutional review board–approved study.

Randomization and warm-up protocol

Individual scheduled surgeries from each recruited surgeon were randomized to either receive warm-up or not, with each surgeon acting as their own control so that intersurgeon performance variability did not impact the ability to observe an intervention effect. The random allocation sequence was generated with random sampling of blocks without replacement and random sampling of surgeries within each block without replacement. Permuted block randomization was conducted, with blocks of 2, 4, and 6 participants each. The random allocation sequence was kept in a REDCap database and revealed one surgery at a time by study staff. Surgeons who were randomized to receive the warm-up curriculum did so on the dVSS for either the da Vinci Si or Xi depending on the

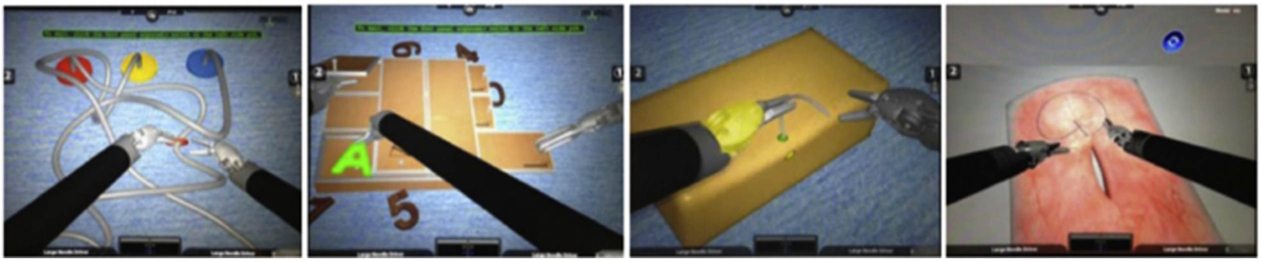

Fig. 2 – Frames from each of the four warm-up modules used during aim 1. From left to right: Ring and Rail, Match Board, Suture Sponge, and Running Suture. (Color version of the figure is available online.)

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

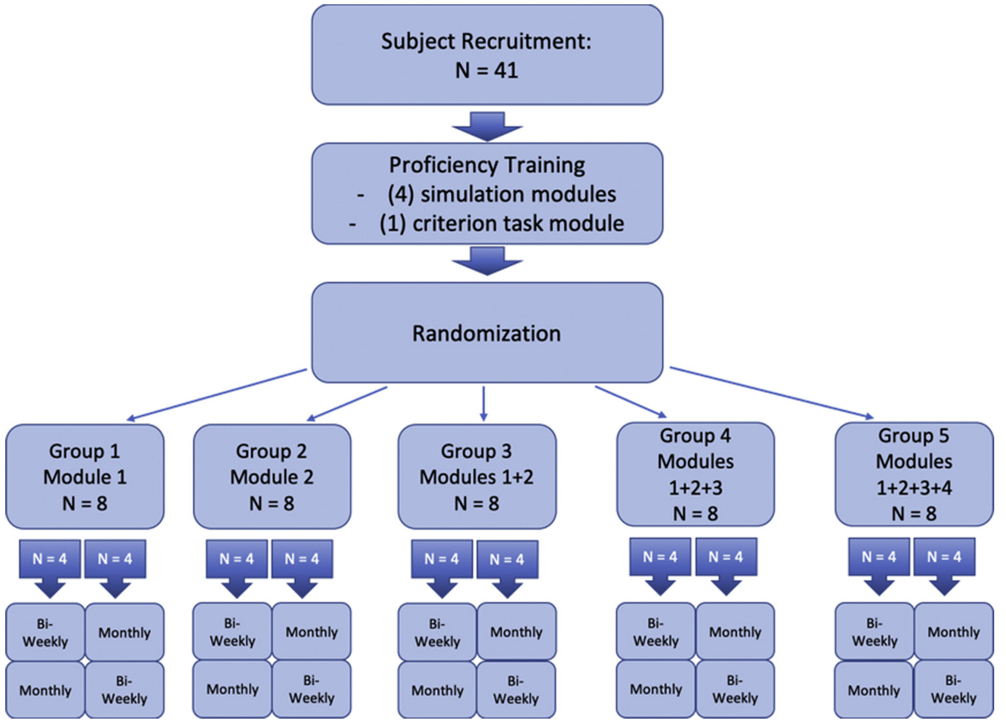

Fig. 3 – Randomization protocol for selection of warm-up module for surgeons. (Color version of the figure is available online.)

available simulator. Participating surgeons completed between 1 and 40 procedures.

Operating room data capture

Endoscopic video and robotic instrument position were captured for each procedure using the Intuitive Surgical dVLogger system which captures data directly from the robot's application programming interface. Capturing these data synchronized to video allowed for assessing performance with the GEARS and tool motion.

Data aggregation and manipulation

All video and tooltip data were captured, and performances were parsed into 'Si' or 'Xi' labels. Any personally identifiable information displayed in the videos was censored to create black boxes over these areas.

Video segmentation

After confirming the timestamps at which the participant surgeon started operating, we obtained the first 15 min of surgery performed. The practicing surgeon's complete video was then uploaded as "unlisted" to YouTube for blinded video review.

Objective metric calculation

The kinematic tooltip data were used to obtain summary metrics based on periods when the surgeon was performing, including ROC, SAL, NAD, MV, laterality, and CUP; all are associated with technical skill.

Video review

For expert surgeon review, a web-based platform was created to distribute the videos accompanied by the GEARS tool to blinded surgeons and nonexpert crowds. Only a subset of randomly chosen videos was used for expert surgeon review as it became unfeasible to have expert surgeons review over 300 videos. Each GEARS score was recorded for each performance to analyze the intraclass correlation (ICC) among ratings for the same video. Statistical calculations were computed using R and Python.[31,32] C-SATS Inc. (Seattle, WA)

Table 1 – The five groups of warm-up modules used for curriculum selection.

Simulation module	Group 1	Group 2	Group 3	Group 4	Group 5
Running Suture 2	X		X	X	X
Ring and Rail 2		X	X	X	X
Match Board 3				X	X
Suture Sponge 3					X

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

was used to compile reviews of technical skill for each video, obtained from crowds of nonexperts, to offset the limited availability of expert surgeons. The use of crowd-sourced assessment of technical skills has been validated as a rapid means of reliably acquiring expert-like quantitative surgical assessments.[33-37]

Results

Aim 1

Forty-one participants completed the study in aim 1, with demographics from these surgeons shown in the supplementary material. As can be seen in Fig. 4, the average TT for the different groups had a downward trend in which the surgeons became quicker as they proceeded through sessions. The "Ring and Rail 2" only module, group 2, was the least effective at lowering TT while also, as seen in Fig. 5, resulted in the lowest GEARS scores. Groups 1 and 5 resulted in the most favorable TT and GEARS scores. Group 1 participants only primed with the "Running Suture 2" module and had a higher GEARS score (2.1 points higher ± 0.86, $P = 0.02$). Compared with group 2, group 1 was on average 31.3 s faster (±14.02 s, $P = 0.03$). GEARS scores from groups 1 and 5 were not significantly different from each other ($P = 0.47$ and $P = 0.70$), although group 1 had the least variation in GEARS scores and TT of all groups. "Running Suture 2", which was also the criterion task, thus was chosen as the warm-up curriculum to use before surgical procedures. It had comparable results as group 5 and practically could be completed in 2-3 min.

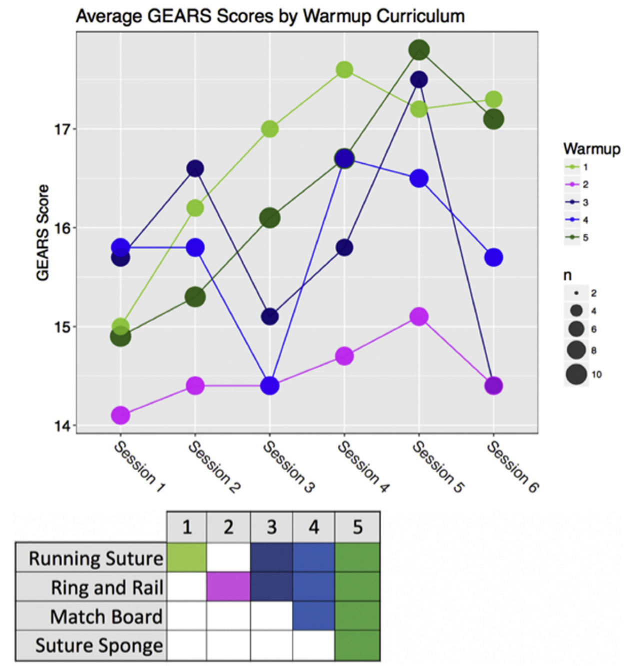

Fig. 5 – The average GEARS score for each of the five warm-up modules in aim 1. Running Suture 2, group 1, had the largest increase in the average GEARS score. (Color version of the figure is available online.)

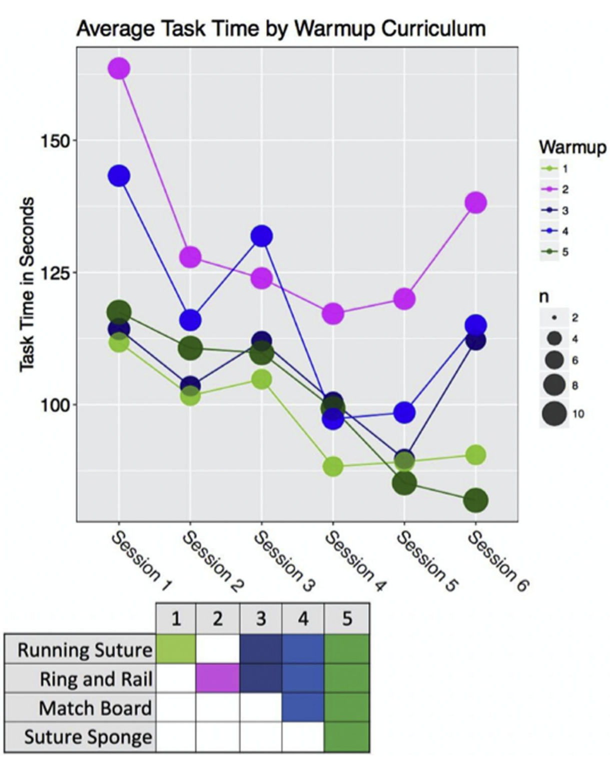

Fig. 4 – Average task time for each of the five warm-up modules in aim 1. Running Suture 2, group 1, had the largest decrease in average task time. (Color version of the figure is available online.)

Aim 2

After identifying the optimal VR warm-up curriculum, 40 practicing surgeons were enrolled in aim 2, with 34 surgeons having performed at least one surgery during the study. Demographic questionnaire results are shown in the supplementary material. Videos of 434 robot-assisted surgeries were performed and recorded. Kinematic data were not captured in 89 cases because of recording equipment technical difficulties, leaving 347 surgeries with both videos and corresponding tooltip position kinematic data.

GEARS review outcome

Ratings by seven expert surgeons were obtained for a subset of 45 videos, with two raters rating all videos and the other five rating from 7 to 44 videos each. ICC for expert ratings was 0.38 (95% confidence interval (CI) 0.25 to 0.54), indicating poor agreement. When the two raters with the greatest numbers of missing observations were removed, ICC was 0.45 (95% CI 0.30 to 0.60), still indicating poor agreement. ICC estimates and their CIs were based on an individual rating, consistency, two-way random-effects model for incomplete data sets.[38]

We assessed the effect of the warm-up on GEARS score, our primary outcome, using both a linear mixed model with a random intercept for each surgeon ($P = 0.98$) and a nonparametric modified Friedman test ($P = 0.34$), with both finding no significant difference between cases receiving or not receiving the warm-up (Table 2). Robot type (Xi *versus* Si) was also not

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

Table 2 – Results from statistical tests on the kinematic and event measures recorded, for MV (mean velocity), NAD (normalized angular displacement), SAL (spectral arc length), CUP (camera use per minute), LR (laterality), and ROC (rate of orientation change).

Measure	Mixed model Est. β (P-value)	Modified Friedman test P-value
MV	0.06 (0.40)	0.35
NAD	−0.06 (0.43)	0.29
SAL	0.06 (0.66)	0.31
CUP	0.03 (0.77)	0.68
LR	−0.01 (0.79)	0.49
ROC	0.02 (0.89)	0.16

Both tests show poor results.

Table 3 – Correlation of the kinematic and event measures to GEARS scores by repeated measured correlation, for MV (mean velocity), NAD (normalized angular displacement), SAL (spectral arc length), CUP (camera use per minute), LR (laterality), and ROC (rate of orientation change).

Measure	Correlation (±95% CI)	P-value
MV	0.08 (0.11)	0.16
ROC	−0.01 (0.11)	0.81
NAD	−0.06 (0.11)	0.32
SAL	−0.02 (0.11)	0.69
LR	−0.01 (0.11)	0.86
CUP	0.08 (0.11)	0.16

These measures were not significantly associated with GEARS scores.

significantly related to the GEARS score ($P = 0.24$), nor was any robot by treatment interaction detected ($P = 0.67$, Fig. 6).

Objective metrics outcome

We also examined the effect of the warm-up on six kinematic metrics MV, ROC, NAD, SAL, CUP, and laterality using both linear mixed models and a nonparametric modified Friedman test, finding no significant differences between treatment arms. We used the Holm-Sidak procedure to maintain familywise alpha at 0.05. Kinematic and event outcomes were not significantly associated with the GEARS score as quantified by repeated measures correlation (using R package rmcorr,[39] results shown in Table 3).

A linear mixed model including both a random intercept and a random slope for the effect of the treatment on each surgeon was used to quantify the range of effects the warm-up had on different surgeons. Treatment effects for individual surgeons on the GEARS score ranged from -0.10 to 0.09 points, which we judged not to be clinically meaningful, thus precluding the possibility of identifying subgroups of surgeons who benefitted from the warm-up to any practically significant degree (Fig. 7).

Discussion

No statistically significant difference in technical skill was observed when surgeons performed a VR warm-up and when they did not. Tests controlling for individual surgeons, robot type (Si *versus* Xi), and time since last surgical performance found no evidence that performing a warm-up curriculum

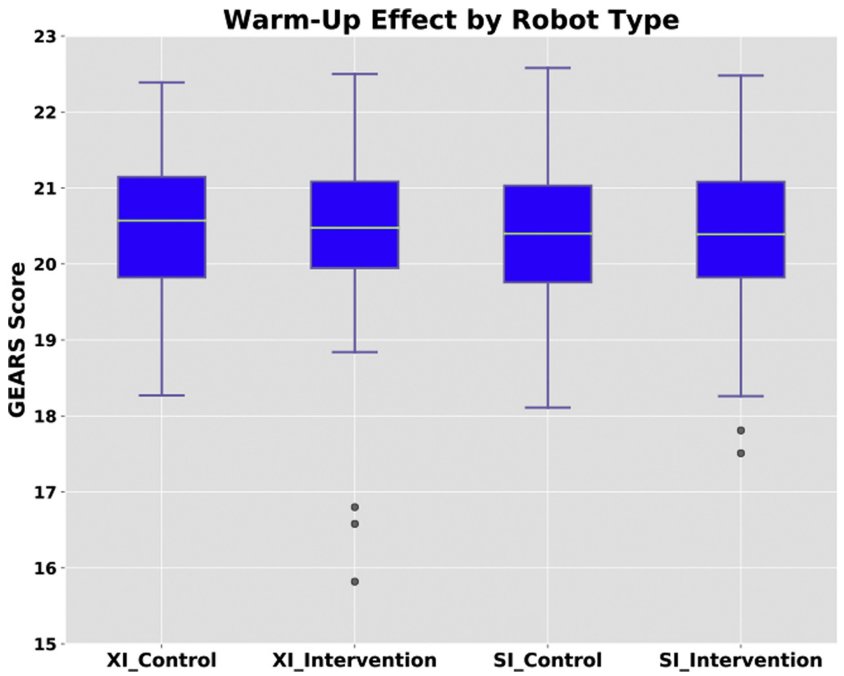

Fig. 6 – Effect of warm-up for each type of the robot used in the study, Xi and Si da Vinci robots. (Color version of the figure is available online.)

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

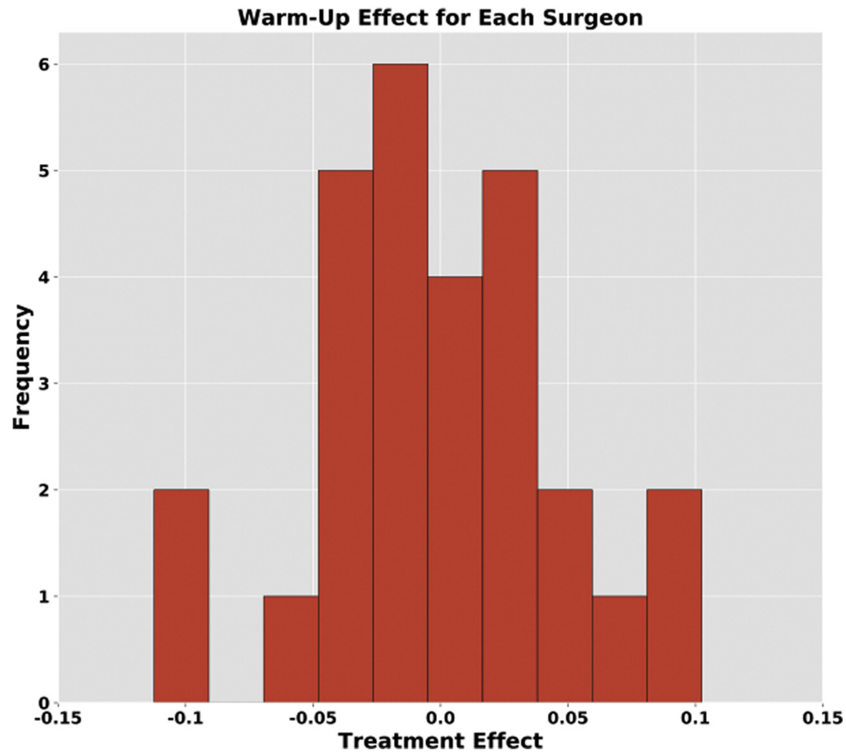

Fig. 7 — A clear normal curve appears from the analysis of warm-up effect for each individual surgeon. This shows strong evidence that the VR warm-up module did not affect technical skill proficiency. (Color version of the figure is available online.)

before surgery elevated surgical readiness. The null findings were unexpected because we had seen a warm-up benefit in our previous dry laboratory study,[19] and the volume of data and trial sessions suggest that the findings from this study are valid.

Sports medicine research concludes that performing a warm-up through the use of proprioceptive neuromuscular facilitation techniques provides improvements to athletic ability and can enhance skill in athletics.[40,41] A study measuring brain function of young athletes performing a warm-up before cycling training found that warm-ups resulted in a significantly increased amount of oxyhemoglobin levels in the prefrontal cortex and motor areas of the brain[42], which is also the case for general motor activity and has been shown to result in shorter reaction times.[43,44] Other studies on motor learning and the brain suggest the prefrontal cortex and anterior cingulate cortex are in control of the initial phases of motor skill learning,[45,46] as well as a study finding behavioral modification to lead to activation in these regions.[47]

These same areas of the brain, related to visual imagery and motion processing, have been found to be activated during surgical performance training tasks,[44,48] suggesting that some form of training or warm-up could potentially aid surgeons in preparing for surgery. A study which included surgeons of mixed experience levels concluded that for less experienced surgeons, a larger modification in prefrontal cortex activation appeared to confer technical performance adaptation, suggesting this warm-up effect may only be present in novice surgeons.[49]

Past studies have shown differing results from the use of a warm-up before surgical procedures. A study performed with residents using laparoscopic procedures found that brief warm-ups performed before surgery significantly improved intraoperative performance, regardless of case difficulty.[50] Alternatively, some studies have shown no warm-up effect. A study of nine laparoscopic surgeons performing no warm-up, laparoscopic trainer box warm-up, or PlayStation 2 warm-up showed that out of 75 recorded cases, warm-up before surgery was not essential to performance, as measured by an objective assessment tool.[51] A randomized, controlled study in which surgeons performed VR warm-up training before laparoscopic cases found that overall performing the warm-up did not increase psychomotor skills during surgery as measured by global assessment scores and also did not influence operating time or complication rates.[52] We previously have shown that in dry laboratory settings of robot-assisted procedures in which surgeons performed VR warm-up before criterion tasks such as suturing, warm-up yielded a significant reduction in errors. Moreover, the study found that expert surgeons experienced a significant improvement from warm-up in their TT and economy of motion, whereas these findings did not extend to less-experienced surgeons.[19]

The several different types of surgeries and surgeons that were included in this study could have been what led to the findings, unless there is indeed no warm-up effect for complex human surgeries. Another possibility is that there are so many actions during the beginning of a surgical case (getting access, cutting skin, etc.) that these hone the attentional capacity of the surgeon such that they are already “warmed-up”, so the VR warm-up is superfluous. Recording the initial 15 min

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

of each surgical case could have also contributed to these findings. Although it is most likely the first 15 min of surgery was the most likely to derive a warm-up benefit, what happens during this time may be relatively simple steps in a case, which would not tax the performer as much as the steps to follow. The VR module used may have not primed surgeons enough to boost arousal or performance to a significant degree. In addition, these modules may stimulate important centers of the brain that confer good surgical performance but more so in the domain of surgical decision-making than of technical skill, and this was not studied.

This research study had limitations. Surgeons performed different types of procedures which consist of several different types of movements and skills needed to complete. Despite all surgeons enrolled being faculty surgeons, varying degrees of skill existed. This study looked at intraoperative performance only and was not designed to assess patient outcomes, thus understanding whether warm-up had an impact on success of surgery could not be ascertained. Assessing these factors was beyond the scope of this study. Warm-up has been shown to have a benefit in conventional laparoscopy intraoperatively,[18,53] and it is possible that robot-assisted surgery may not be as technically challenging, thus not requiring a technical skill warm-up. Preoperative warm-up may take the form of mental visualization, reviewing patient images beforehand, describing steps of a case to a trainee, patient positioning, placement of robotic ports, and initial conventional laparoscopic tissue manipulation. These are all important and may have a significant positive impact on a surgeon's priming such that assessment of technical skills alone may not have detected the added value of VR warm-up.

Conclusions

Counter to most existing literature in conventional laparoscopic surgery and observations made by this author group in a robot-assisted surgery dry laboratory setting, performing a VR warm-up curriculum before robot-assisted human surgery does not significantly improve the technical skill of a practicing surgeon. Because surgery is a dynamic process requiring thoughtful choreography before, during, and after a case, isolating a specific preoperative surgeon behavior or effort that may impact the surgeon's readiness may be difficult.

Acknowledgment

We thank each of the faculty surgeons for giving their time to provide us with video review for several videos of robotic surgery procedures: Ardavan Akhavan, Walid Farhat, Christopher Weight, Jonathan Ellison, Tim Brand, Thomas Lendvay, and Matthew Sorensen. We are grateful for Drs. Simon DiMaio, Omid Mohareri, Anthony Jarc, and Ms. Cortney Jansen from Intuitive Surgical who provided appreciated technical support during the data collection phase of the project and helped provide access to video and synchronized robot motion data. In addition, we would like to thank all the surgeons and residents who graciously participated in this study and committed so much of their time.

Authors' contributions

T.S.L., M.D.S., R.S., T.M., and T.M.K. contributed to study conception and design. L.M. contributed to acquisition of data. J.D.K., A.F., M.N., N.H., N.O., E.G., B.C., and T.M.K. contributed to analysis and interpretation of data. J.D.K., L.M., M.N., T.S.L., and T.M.K. contributed to drafting of the manuscript.

Funding sources

This study was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-15-2-0030. The views, opinions, and endorsement by the author(s) do not reflect those of the US Army or the Department of Defense.

The Seattle Children's Core for Biomedical Statistics is supported by the Center for Clinical and Translational Research at Seattle Children's Research Institute and grant UL1RR025014 from the NIH National Center for Research Resources.

Disclosure

The authors would like to disclose the following potential conflict of interest: T.S.L. was an equity owner of CSATS, Inc., until September 19, 2019, and has a patent regarding crowd-sourced assessment of performance of an activity, licensed to Johnson and Johnson, Inc. The other authors report no proprietary or commercial interest in any product mentioned or concept discussed in this article.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jss.2021.01.037.

REFERENCES

1. Kohn L, JM C, Donaldson M, eds. *To Err is Human: Building a Safer Health Care System*. Washington, DC: National Academy Press; 2000.
2. Zhan C, Miller M. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA*. 2003;290:1868.
3. Murphy S, Xu J, Kochanek K. Deaths: preliminary data for 2010. *Natl Vital Stat Rep*. 2012;60; 2012. https://www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_04.pdf. Accessed May 1, 2020.
4. Shreve J, , Actuaries Health Section S, Firm M. The economic measurement of medical errors. Society of Actuaries. 2010. Available at: http://soa.org/Files/Research/Projects/research-econ-measurement.pdf.
5. Sroka G, Feldman L, Vassiliou M, et al. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room - a randomized controlled trial. *Am J Surg*. 2010;199:115.
6. Stefanidis D, Korndorffer Jr JR, Sierra R, Touchard C, Dunne JB, Scott DJ. Skill retention following proficiency-based laparoscopic simulator training. *Surgery*. 2005;138:165–170.
7. Perez RS, Skinner A, Weyhrauch P, et al. Prevention of surgical skill decay. *Mil Med*. 2013;178:588–592.

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

8. Sheetz KH, Claflin J, Dimick JB. Trends in the adoption of robotic surgery for common surgical procedures. JAMA Netw Open 3: 2020.
9. Haynes AB, Weiser TG, Berry WR, et al. A Surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med*. 2009;360:491–499.
10. Nascon J, Schmidt RA. The activity-set hypothesis for warm-up decrement. *J Mot Behav*. 1971;3:1–16.
11. Wrisberg CA, Salmoni AW, Schmidt RA. Warm-up effects in the learning of discrete motor skills. *Acta Psychol*. 1975;29:311–320.
12. Anshel MA. The effect of arousal on warm-up decrement. *Res Q Exerc Sport*. 1985;56:1–9.
13. Anshel MH, Wrisberg CA. Reducing warm-up decrement in the performance of the tennis serve. *J Sport Exerc Psychol*. 1993;15:290–303.
14. Wetter PA. *Resubmission of gap analysis workshop for training for reintegration of surgical skills*. Miami, FL: Proceedings of the 2011 Society of Laparoendoscopic Surgeons Conference; 2011.
15. Arthur Jr W, Winston Jr B, Stanush PL, McNelly TL. Factors that influence skill decay and retention: a quantitative review and analysis. *Hum Perform*. 1998;11:57–101.
16. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg*. 2002;236:458–463.
17. Kahol K, Satava R, Ferrara J, Smith ML. Effect of short-term pretrial practice on surgical proficiency in simulated environments: a randomized trial of the "preoperative warm-up" effect. *J Am Coll Surg*. 2009;208:255.
18. Calatayud D, Arora S, Aggarwal R, et al. Warm-up in a virtual reality environment improves performance in the operating room. *Ann Surg*. 2010;251:1181.
19. Lendvay TS, Brand TC, White L, et al. Virtual reality robotics surgery warm-up improves surgical performance: a prospective randomized controlled study. *J Am Coll Surg*. 2013;216:1181–1192.
20. Stefanidis D, Korndorffer Jr JR, Sierra R, Touchard C, Dunne JB, Scott DJ. Skill retention following proficiency-based laparoscopic simulator training. *Surgery*. 2005;138:165–170.
21. Mimic Technologies Inc. *Da Vinci skills simulator (dvss)*; 2018. Mimic Technologies Inc.; 2018. https://mimicsimulation.com/. Accessed May 1, 2020.
22. Holst D, Kowalewski T, White L, et al. Crowd-sourced assessment of technical skills: an adjunct to urology resident surgical simulation training. *J Endourol*. 2015;29:604–609.
23. Kowalewski T, Comstock B, Sweet R, et al. Crowd-sources assessment of technical skills for validation of basic laparoscopic urologic skills tasks. *J Urol*. 2016;195:1859–1865.
24. Chen J, Oh PJ, Chang N, et al. Use of automated performance metrics to measure surgeon performance during robotic vesicourethral anastomosis and methodical development of a training tutorial. *J Urol*. 2018;200:895–902.
25. Nguyen JH, Chen J, Mashall SP, et al. Using objective robotic automated performance metrics and task-evoked pupillary response to distinguish surgeon expertise. *World J Urol*. 2019.
26. Balasbramanian S, Melendez-Calderon A, Burdet E. A robust and sensitive metric for quantifying movement smoothness. *IEEE Trans Biomed Eng*. 2012;59.
27. Kowalewski TM, White LW, Lendvay TS, et al. Beyond task time: automated measurements augments fundamentals of laparoscopic skills methodology. *J Surg Res*. 2014;192:329–338.
28. Sharon Y, Lendvay TS, Nisky I. Instrument orientation metrics for robot-assisted and open surgical skill evaluation. *IEEE Trans Hum Mach Syst*. 2017.
29. Hung AJ, Zehnder P, Patil MB, et al. Face, content and construct validity of a novel robotic surgery simulator. *J Urol*. 2011;186:1019–1025.
30. Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J Urol*. 2012;187:247–252.
31. Python Software Foundation. Python Language Reference. Available at: http://www.python.org. Accessed May 1, 2020.
32. R Core Team. R: a language and environment for statistical computing, Vienna, Austria. 2018. Available at: http://www.R-project.org. Accessed May 1, 2020.
33. Chen SP, Kirsch S, Zlatev DV, et al. Optical biopsy of bladder cancer using crowd-sourced assessment. *JAMA Surg*. 2016;151:90–93.
34. Lendvay TS, White L, Kowalewski T. Crowdsourcing to assess surgical skill. *JAMA Surg*. 2015;150:1086–1087.
35. Chen SP, Kirsch S, Zlatev DV, et al. Optical biopsy of bladder cancer using crowd-sourced assessment. *JAMA Surg*. 2016;151:90–93.
36. Holst D, White L, Brand T, Harper J, Lendvay T. *Crowd-sourced assessment of technical skills (C-SATS): an adjunct to urology resident simulation training*. Orlando, FL: American Urological Association Annual Meeting; 2014.
37. White L, Kowalewski T, Brand T, et al. *Crowd-sourced assessment of technical skills (C-SATS): fast, economical, and accurate assessment of robotic surgery*. Orlando, FL: American Urological Association Annual Meeting; 2014.
38. Coefficients of interrater reliability - generalized for randomly incomplete datasets. Brueckl M and Heuer F. 2018. Available at: https://cran.r-project.org/web/packages/irrNA/index.html. Accessed May 1, 2020.
39. Repeated measures correlation. Jonathan Z. Bakdash and Laura R. Marusich. 2017. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5383908/. Accessed May 1, 2020.
40. Shellock FG, Prentice WE. Warming-up and stretching for improved physical performance and prevention of sports-related injuries. *Sports Med*. 1985;2:267–278.
41. Daneshjoo A, Mokhtar AH, Rahnama N, Yusof A. Effects of the 11+ and harmoknee warm-up programs on physical performance measures in professional soccer players. *J Sports Sci Med*. 2013;12:489–496.
42. Techayusukcharoen R, Iida S, Aoki C. Observing brain function via functional near-infrared spectroscopy during cognitive program training (dual task) in young people. *J Phys Ther Sci*. 2019;31.
43. Maki A, Yamashita Y, Ito Y, Watanabe E, Mayanagi Y, Koizumi H. Spatial and temporal analysis of human motor activity using noninvasive NIR topography. *Med Phys*. 1995;22:1997–2005.
44. Leff DR, Orihuela-Espina F, Atallah L, et al. Functional prefrontal reorganization accompanies learning-associated refinements in surgery: a manifold embedding approach. *Comput Aided Surg*. 2008;13:325–339.
45. Kelly AM, Garavan H. Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex*. 2005;15:1089–1102.
46. Halsband U, Lange RK. Motor learning in man: a review of functional and clinical studies. *J Physiol Paris*. 2006;99:414–424.
47. Walton MR, Croxson PL, Behrens TEJ, Kennerley SW, and Rushworth MFS. Adaptive decision making and value in the anterior cingulate cortex. Neuroimage 36: 2007.
48. Wanzel KR, Anastakis DJ, McAndrews MP, et al. Visual-spatial ability and fMRI cortical activation in surgery residents. *Am J Surg*. 2007;193:507–510.
49. Leff DR, Elwell CE, Orihuela-Espina F, et al. Changes in prefrontal cortical behaviour depend upon familiarity on a

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.

bimanual coordination task: an fNIRS study. *Neuroimage*. 2008;39:805–813.
50. Chen CCG, Green IC, Colbert-Getz Jm, et al. Warm-up on a simulator improves residents' performance in laparoscopic surgery: a randomized trial. *Int Urogynecol J*. 2013;24:1615–1622.
51. Weston MK, Stephens JH, Schafer A, Hewett PJ. Warm-up before laparoscopic surgery is not essential. *Anz J Surg*. 2014;84:143–147.
52. Polterauer S, Husslein H, Kranawetter m, et al. Effect of preoperative warm-up exercise before laparoscopic gynecological surgery: a randomized trial. *J Surg Educ*. 2016;73:429–432.
53. Lee JY, Mucksavage P, Kerbl DC, et al. Laparoscopic warm-up exercises improve performance of senior-level trainees during laparoscopic renal surgery. *J Endourol*. 2012;26:545–550.

Downloaded for Anonymous User (n/a) at University of Minnesota Twin Cities from ClinicalKey.com by Elsevier on August 15, 2023. For personal use only. No other uses without permission. Copyright ©2023. Elsevier Inc. All rights reserved.