STUDY PROTOCOL

Validity Evidence for Simulation-based Assessment of Robotic Cardiac Surgical Skills and Examination of Learning Curves in Wet Labs

February 11, 2021

ClinicalTrial.gov NCT05043064

Study rationale

Robotics has entered the mainstream of cardiac surgery, with large numbers of cases now performed at specialized centers[1,2,3]. But despite the demonstrated benefits and increasing number of robotic cases, there has been slow adoption of this technology partly due to lack of proper training[4]. An important aspect of modern surgical education is that it utilizes competency-based training instead of time-based training or numbers of performed procedures[5]. The classic stepwise educational approach in robotic cardiac surgery involves increasing entrustment with increasing levels of surgical skills[6]. Until now, this approach has relied on clinical training with little use of simulation-based training as initial part of the entrustment process[7]. Simulation-based training allows for skill development outside of the operating room in a controlled and safe environment with no patient risk and ample opportunities for feedback and correction of errors[8].

A recent study[7] compared currently available simulation modalities used in cardiac robotic surgical training (wet lab, dry lab and virtual reality) and indicated that trainees had the greatest skill improvement in the wet lab. However, the study was not designed to examine validity evidence of the wet lab model and had a small number of participants (n=10) in the wet lab group.

The skills and performance levels of trainees need to be objectively assessed to ensure the surgeons competency level before they operate on real patients[9]. Before using simulation-based training to assess competency it is imperative that the assessment tool is examined to ensure it is supported by validity evidence[10,11]. Currently, it is unclear how much wet lab robotic cardiac surgery simulation training is needed before a trainee can proceed to real operations[12] and there is a need for identifying a relevant pass/fail level to ensure competency before operating on patients[11].

Learning curves are increasingly used in the assessment of competency and training program design[13,14,15]. Methods currently used to analyze learning curves in surgery are mainly descriptive and need to be rigorous with quantifiable parameters to allow future surgeons to benefit from established performance standards[13]. Data on learning curves in robotic cardiac procedures in a real-world setting exists[16,17], but little is known about learning curves in a wet lab simulation setting. Learning curves can support a competence-based approach to assessment for learning[14]. It has been claimed that surgeon expertise in performing both routine and high-risk cardiac procedures before proceeding to robotic-assisted surgery is essential[4]. However, it is possible that as robotic simulators evolve, the non-robotic experience gap may be substantially reduced[16]. By contrast, according to the expertise reversal effect, an instructional design that is beneficial to a novice learner may be detrimental to a more experienced learner and vice versa[18]. The skills acquisition rate of robotic novices with expertise in cardiac surgery may be faster than robotic novices with limited cardiac surgery experience, which may suggest that robotic skills training should only be commenced after the surgeons master the non-robotic techniques. However, if this is not the case, robotic skills training programs should be implemented early in cardiac surgeons' careers to allow them to develop expert-level skills as early in their careers as possible.

The current study will be the first to examine validity evidence for the wet lab robotic cardiac surgery model with the largest number of participants to-date and will be the first to analyze learning curves in wet lab simulation settings for subjects with different surgical background. Results of the study could help in establishing standardized training pathways in robotic cardiac surgery as recommended at the ORSI Consensus Meeting on European Robotic Training (OCERT)[7] and facilitate faster spread of robotic cardiac surgery. Most importantly, establishing valid assessment methods will allow us to ensure that robotic surgeons are trained until proficiency in the wet lab before operating being allowed to progress on real patients. This will ensure that all patients are certain to receive expert-level care regardless of where and by whom, they are treated.

Objectives

Primary:

- to investigate validity evidence for simulation-based assessment using the wet lab robotic cardiac surgery model.

Secondary:

- to analyze performance curves in the wet lab for a robotic cardiac surgery model.
- to assess rate of skills acquisition of cardiac robotic skills in cardiac surgery trainees versus non-cardiac surgery trainees.

Materials and methods

We aim to assess the validity evidence for the wet lab robotic cardiac surgery simulation model using the contemporary framework for validity proposed by Messick[147], which includes five sources of validity evidence: content evidence, response process, relations to other variables, internal structure, and consequences.

We will test the validity evidence for the three wet lab tasks used in robotic cardiac surgery using three groups of participants: experienced robotic cardiac surgeons (5), robotic cardiac surgery novices (8) and robotic non-cardiac surgery novices (8).

Wet lab simulator: tasks will be performed in porcine models with the da Vinci Xi Surgical System (Intuitive Surgical, California, USA).

1. Content evidence:

Three wet lab tasks, which all are parts of real robotic cardiac operations have been chosen by surveying experienced robotic cardiac surgeons participating in the study:

- robotic harvesting of 10 cm internal thoracic artery (ITA) from the porcine chest wall
- robotic-assisted placement of 5 sutures in the mitral annulus of the porcine heart (one on each trigone and three consecutive annular sutures)
- robotic-assisted closure of the porcine atrium

2. Response process:

All participants will receive a short oral introduction to the da Vinci Xi Surgical System and will be shown videos of robotically assisted cardiac procedures highlighting basic operative techniques and relevant anatomy. They will then complete the three previously selected tasks to familiarize themselves with the wet lab simulator. Finally, they will complete all 3 tasks again and this will be used to assess their base-line performance.

Next, five experienced robotic cardiac surgeons from different international sites will perform each task 4 times more. Subjects in the cardiac and non-cardiac groups will perform the tasks until mastery is achieved based on the performance of the expert robotic surgeons. To ensure the successful completion of the task, each participant will be required to pass each exercise 2 consecutive times with the only one failure between the attempts allowed.

The mastery learning level for the tasks will be calculated from the mean time-based score (TBS) of the experienced surgeons 2 last attempts for each task.
Time-based scoring equation will be calculated as follows (adapted from Valdis et al[6]):

Score = Max time – Completion time – Errors

-Max time is the total time that an individual will be allowed to complete the task during the base-line performance.
-Completion time is either;

1) The mean time for completion of each expert's last 2 attempts which will be used to calculate the mastery learning level based on the formula above
2) The time taken by each study participant to complete the task

- Errors will be divided in minor and major ones. A minor error will result in adding a 10 second penalty time to the TBS equation. A major error will result in a score of zero, regardless of time to completion.

Minor errors:

1) moving robotic arms out of view
2) robotic arm collision
3) drop of the needle

Major errors:

1) avulsion of ITA or suturing model
2) tearing/ fraying the suture
3) gross tissue damage (specifically with the cautery)

All tasks will be recorded on the robot's camera and evaluated blinded using the modified Global Evaluative Assessment of Robotic Skills (mGEARS) score (Appendix 1) by two not participating surgeons with experience in robotic cardiac surgery from two different sites in Europe and North America. In modified GEARS score the tasks will not be rated on autonomy domain as autonomy cannot be evaluated from recordings.

3. Relations to other variables:

We will analyze variation in performance scores (TBS and mGEARS) between experienced robotic cardiac surgeons and robotic novices during their base-line performance. We will use TBS score for the assessment of novice performance during the subsequent training rounds.

4. Internal structure:

To examine the reliability of the final simulator task, the internal consistency and test/retest reliability will be assessed. The internal consistency of TBS and mGEARS scores will be used for the assessment of trainee performance for the base-line task. The test/retest reliability will be calculated by comparing the trainees' TBS and mGEARS performance in the last 2 training sessions.

5. Consequences:

To investigate the consequence of testing we will define a pass/fail level to be used for mastery learning. The pass/fail level will be determined using the “contrasting groups method” for the mean TBS and mGEARS score for the last two attempts in the experienced group.

Selection of participants:

Inclusion criteria

- Experienced robotic cardiac surgeons: robotic surgeons with total volume > 50 robotic cardiac operations and at least 20 robotic cases annually[4].
- Cardiac surgery novices: on-going or completed cardio-thoracic residency.
- Non-cardiac surgery novices: on-going or completed surgical specialty residency (gynecology, urology, general surgery, ENT surgery).

Exclusion criteria

- Novices with > 5 hours experience of any robotic system.

Outcomes:

Primary outcome measures:

- time-based score for each task

Secondary outcome measures:

- modified Global evaluative assessment of robotic skills score for each task
- the training time to reach the mastery learning level for the novice group and cardiac surgery novices; both total time and for each task.

Statistical plan and data analysis

Within the current literature exists the only study assessing robotic cardiac surgery simulation training with sample size of 10 participants in the wet lab group. The reports on growth models and learning curve dataset recommend least 3 repeated observations per individual[15]. We plan to recruit 16 participants in the novice and 5 robotic cardiac surgeons from different international sites in the experienced group. We plan to include equal numbers of cardiac (8) and non-cardiac trainees (8). Additionally, two surgeons with experience in robotic cardiac surgery from two different sites will be recruited to rate all the training sessions.

Based on the previous study[7] we hypothesize that the subjects in the novice group will be able to achieve the mastery learning within 5 trainings sessions and thus fulfil the requirement of at least 3 repeated observations.

STATA will be used for statistical analysis. The Kruskal-Wallis test will be used to examine a difference between the novice and the expert group at base-line. We will use the Mann-Whitney U test to compare continuous variables. A two-sided significance level of 0.05 will be used. The internal consistency will be evaluated by Cronbach’s alpha, and the test/retest reliability of the last two training sessions using intra-class correlation coefficient.

APPENDIX 1. Modified global evaluative assessment of robotic skills (mGEARS)

Depth perception

1	2	3	4	5
Constantly overshoots target, wide swings, slow to correct		Some overshooting or missing target, but quick to correct		Accurately directs instruments in the correct plane to target

Bimanual dexterity

1	2	3	4	5
Uses only one hand, ignores nondominant hand, poor and coordination		Uses both hands, but does no optimize interaction between hands		Expertly uses both hands in a complementary way to provide best exposure

Efficiency

1	2	3	4	5
Inefficient efforts; many uncertain movements, constantly changing focus or persisting without progress		Slow, but planned movement are reasonably organized		Confident, efficient and safe conduct, maintains focus on task, fluid progression

Force sensitivity

1	2	3	4	5
Rough moves, tears tissue, injures nearby structures, poor control, frequent suture breakage		Handles tissues reasonably well, minor trauma to adjacent tissue, rare suture breakage		Applies appropriate tension, negligible injury to adjacent structures, no suture breakage

Robotic control

1	2	3	4	5
Consistently does not optimize view, hand position, or repeated collisions even with guidance		View is sometime not optimal. Occasionally needs to relocate arms. Occasional collisions and obstruction of assistant.		Controls camera and hand position optimally and independently. Minimal collisions or obstruction of assistant.

References

1. Gillinov AM, Mihaljevic T, Javadikasgari H, et al. Early results of robotically assisted mitral valve surgery: Analysis of the first 1000 cases. *J Thorac Cardiovasc Surg*. 2018;155(1):82-91.e2. doi:10.1016/j.jtcvs.2017.07.037
2. Bonaros N, Schachner T, Lehr E, et al. Five hundred cases of robotic totally endoscopic coronary artery bypass grafting: Predictors of success and safety. *Ann Thorac Surg*. 2013;95(3):803-812. doi:10.1016/j.athoracsur.2012.09.071
3. Pettinari M, Navarra E, Noirhomme P, Gutermann H. The state of robotic cardiac surgery in Europe. *Ann Cardiothorac Surg*. 2017;6(1):1-8. doi:10.21037/acs.2017.01.02
4. Rodriguez E, Nifong LW, Bonatti J, et al. Pathway for surgeons and programs to establish and maintain a successful robot-assisted adult cardiac surgery program. *J Thorac Cardiovasc Surg*. 2016;152(1):9-13. doi:10.1016/j.jtcvs.2016.05.018
5. Havemann MC, Dalsgaard T, Sørensen JL, et al. Examining validity evidence for a simulation-based assessment tool for basic robotic surgical skills. *J Robot Surg*. 2019;13(1):99-106. doi:10.1007/s11701-018-0811-8
6. Schachner T, Bonaros N, Wiedemann D, et al. Training Surgeons to Perform Robotically Assisted Totally Endoscopic Coronary Surgery. *Ann Thorac Surg*. 2009;88(2):523-527. doi:10.1016/j.athoracsur.2009.04.089
7. Valdis M, Chu MWA, Schlachta C, Kiaii B. Evaluation of robotic cardiac surgery simulation training: A randomized controlled trial. *J Thorac Cardiovasc Surg*. 2016;151(6):1498-1505.e2. doi:10.1016/j.jtcvs.2016.02.016
8. Cook DA, Hatala R, Brydges R, et al. Technology-Enhanced Simulation for Health Professions Education. *JAMA*. 2011;306(9). doi:10.1001/jama.2011.1234
9. Vanlander AE, Mazzone E, Collins JW, et al. Orsi Consensus Meeting on European Robotic Training (OCERT): Results from the First Multispecialty Consensus Meeting on Training in Robot-assisted Surgery. *Eur Urol*. 2020:1118. doi:10.1016/j.eururo.2020.02.003
10. Downing SM. Validity: On the meaningful interpretation of assessment data. *Med Educ*. 2003;37(9):830-837. doi:10.1046/j.1365-2923.2003.01594.x
11. Jarocki A, Rice D, Kent M, Oh D, Lin J, Reddy RM. Validity of robotic simulation for high-stakes examination: a pilot study. *J Robot Surg*. 2021;(0123456789). doi:10.1007/s11701-021-01258-9
12. Gillinov M, Mick S, Mihaljevic T, Suri RM. Watch one, do one, teach one. *J Thorac Cardiovasc Surg*. 2016;151(6):1506-1507. doi:10.1016/j.jtcvs.2016.02.033
13. Valsamis EM, Chouari T, O'Dowd-Booth C, Rogers B, Ricketts D. Learning curves in surgery: Variables, analysis and applications. *Postgrad Med J*. 2018;94(1115):525-530. doi:10.1136/postgradmedj-2018-135880
14. Pusic M V., Boutis K, Hatala R, Cook DA. Learning Curves in Health Professions Education. *Acad Med*. 2015;90(8):1034-1042. doi:10.1097/ACM.0000000000000681
15. Pusic M V., Boutis K, Pecaric MR, Savenkov O, Beckstead JW, Jaber MY. A primer on the statistical modelling of learning curves in health professions education. *Adv Heal Sci Educ*. 2017;22(3):741-759. doi:10.1007/s10459-016-9709-2
16. Goodman A, Koprivanac M, Kelava M, et al. Robotic Mitral Valve Repair: The Learning Curve. *Innov Technol Tech Cardiothorac Vasc Surg*. 2017;12(6):390-397. doi:10.1097/imi.0000000000000438
17. Yaffee DW, Loulmet DF, Kelly LA, et al. *Can the Learning Curve of Totally Endoscopic Robotic Mitral Valve Repair Be Short-Circuited?* Vol 9.
18. Kalyuga S, Ayres P, Chandler P, Sweller J. The expertise reversal effect. *Educ Psychol*. 2003;38(1):23-31. doi:10.1207/S15326985EP3801_4