DOI:
10.1007/s00464-017-5622-x

Development and Validation of a Tool for Non-Technical Skills Evaluation in Robotic Surgery- The ICARS System

Nicholas Raison MRCS[1,2], Thomas Wood BSc[1], Oliver Brunckhorst MBBS[1], Takashige Abe PhD[1], Talisa Ross BSc[1], Ben Challacombe FRCS[3], Mohammed Shamim Khan FRCS[1], Giacomo Novara MD[4], Nicolo Buffi MD[5], Henk Van Der Poel PhD[6], Craig Mcilhenny FRCS[7], Prokar Dasgupta MD[1], Kamran Ahmed PhD[1]

Affiliation:
1. King's College London, MRC Centre for Transplantation, London, UK
2. The London Clinic, Department of Urology, London, UK
3. Guy’s and St Thomas’ NHS Trust, Department of Urology, London, UK
4. University of Padua, Department of Urology, Padua, Italy
5. Humanitas Clinical and Research Centre, Department of Urology, Rozzano, Milan, Italy
6. Netherlands Cancer Institute, Amsterdam, The Netherlands
7. NHS Forth Valley, Department of Urology, Larbert, United Kingdom

Corresponding Author

Prof P Dasgupta
Chair in Robotic Surgery & Urological Innovation, Honorary Consultant Urological Surgeon
Division of Transplantation Immunology & Mucosal Biology
MRC Centre for Transplantation
King's College London
Faculty of Life Sciences & Medicine
5th Floor Tower Wing,
Guy's Hospital, London SE1 9RT
Email: prokar.dasgupta@kcl.ac.uk
Tel: 0207 188 5906
Fax: 0207 188 0944

Abstract

Background

Non-technical skills (NTS) are being increasingly recognised as vital for safe surgical practice. Numerous NTS rating systems have been developed to support effective training and assessment. Yet despite the additional challenges posed by robotic surgery, no NTS rating systems have been developed for this unique surgical environment. This study reports the development and validation of the first NTS behavioural rating system for robotic surgery.

Methods

A comprehensive index of all relevant NTS behaviours in robotic surgery was developed through observation of robotic theatre and interviews with robotic surgeons. Using a Delphi methodology, a panel of 16 expert surgeons were consulted to identify behaviours important to NTS assessment. These behaviours were organised into an appropriate assessment template. Experts were consulted on the feasibility, applicability and educational impact of ICARS.

An observational trial was used to validate ICARS. 73 novice, intermediate and expert robotic surgeons completed a urethrovesical anastomosis within a simulated operating room. NTS were tested using four scripted scenarios of increasing difficulty. Performances were video recorded. Robotic and NTS experts assessed the videos post-hoc using ICARS and the standard behavioural rating system, NOn-Technical Skills for Surgeons (NOTSS).

Results

28 key non-technical behaviours were identified by the expert panel. The finalised behavioural rating system was organised into 4 principle domains and 7 categories. Expert opinion strongly supported its implementation. ICARS was found to be equivalent to NOTSS on Bland-Altman analysis and accurately differentiated between novice, intermediate and expert participants, p=0.01. Moderate agreement was found between raters, Krippendorff's alpha = 0.4. The internal structure of ICARS was shown to be consistent and reliable (median Cronbach alpha = 0.92, range 0.85-0.94).

Conclusion

ICARS is the first NTS behavioural rating system developed for robotic surgery. Initial validation has shown it to be an effective and reliable tool. Implementation of ICARS will supported structured training and assessment of NTS within the robotic surgical curriculum.

Keyword: **Robotic surgery; non-technical skills; education; curriculum; assessment tool; training**

Introduction

Effective non-technical skills (NTS) training and assessment is becoming an increasingly important component of the medical curriculum. The growing recognition of the role of NTS is clearly reflected in the quality and quantity of the behavioural training tools now available[1].

Provision of safe and reliable patient has always been of paramount importance in surgery, but this has been matched in recent years by large expansions in the regulatory monitoring of outcomes. Such scrutiny has helped to highlight that that technical competency alone does not guarantee success[2]. Recognition of the vital importance of NTS has resulted in the development of various NTS behavioural rating systems. Separate systems have been developed for assessment of the entire surgical team[3-5] as well as individual team members such as surgeons[6,7], anaesthetists[8] and scrub practitioners[9]. To be effective such rating systems must accurately capture relevant NTS behaviours. Established systems such as Non-Technical Skills for Surgeons (NOTSS) taxonomy have been applied to a variety of surgical specialties but in highly specialist environments, such generic tools may not be suitable[10].

Robotic surgery has expanded rapidly in recent years despite high costs limiting its use to specialist centres. Given the limited availability of robotic systems, training remains challenging and various simulation based training curricula have been developed[11]. Robotic surgery demands significant adaptations to the standard operating room (OR) environment including team interaction. As a result, proficiency in robotic surgery requires specialist training in both technical and non-technical skills. Despite this, only recently have robotic surgical curricula begun to incorporate NTS [12,13] and no behavioural markers systems have yet been developed for robotic surgery[1].

This study aims to develop the first behavioural rating system specifically for surgeons NTS during robotic surgery and provide initial validation evidence to demonstrate its applicability to this unique operating environment.

Materials and Methods

Development of the Interpersonal and Cognitive Assessment for Robotic Surgery (ICARS) Behavioural Rating System

This prospective study was conducted in collaboration with the Urological Department, Guy's Hospital, the Vattikuti Robotic Surgery Training Centre and the MRC Centre for Transplantation, King's College London. Approval was granted by the local research ethics committee (Study Reference: LRU15/162080).

ICARS was developed using a similar core framework to those described by previous role specific behavioural rating systems [6,9,8]. Development of ICARS was performed in three stages. In the first stage, a full taxonomy of NTS behaviours relevant to surgeons performing robotic surgery was compiled. Secondly the key NTS behaviours in robotic surgery were identified using a modified Delphi process involving a panel of expert surgeons. In the final stage, a behavioural marker system was designed, incorporating the key NTS skills for robotic surgery into a practical checklist. The checklist was reviewed by the consensus panel for final approval.

To create a comprehensive record of non-technical skills, 15 hours of live robotic surgery were observed by two authors trained in surgical NTS. All potential behaviours and skills relevant to non-technical skills assessment were recorded and collated. Throughout this process, interviews were conducted with the operating surgeons to gain further insight into potentially relevant NTS behaviours which were then correlated with clinical observations. This provided a comprehensive catalogue of non-technical skills in robotic surgery. The final list was decided by agreement of the researchers before being grouped and sub-categorised into a preliminary checklist.

The draft ICARS rating system was disseminated to a consensus panel consisting of 16 surgeons (10 expert robotic surgeons and 6 expert laparoscopic surgeons with an intermediate proficiency in robotic surgery). Surgeons were recruited to the panel by invitation on the basis of their experience. A Delphi process was used to refine the checklist and identify the key relevant behaviours for NTS assessment. In the first round, the comprehensive checklist of behaviours was distributed. Panel members were asked to rate the importance of each component for checklist inclusion using a 5-point Likert scale (5-definitely important to 1- definitely not important). Inclusion was determined by a mean score of ≥4. Excluded components were removed and the Delphi process continued until there was saturation of information.

The approved list was then categorised and formatted into an appropriate assessment template. A five-point rating scale was chosen to rate the specific behaviours matched to subjective standards to aid assessment. These markers ranged from 1 = unacceptable to 5 = excellent. This scale was chosen to ensure raters had specific scope for differentiating subjects.

The finalised ICARS rating system was then recirculated amongst the experts. All members of the consensus panel approved the finalised ICARS rating system. Expert opinion on the assessment tool was formally gathered using a qualitative and quantitative questionnaire.

Validation Protocol

An observational trial was undertaken to validate the ICARS behavioural rating system. Participants were invited to take part in the study as part of a larger trial of cognitive training in surgery (trial reference ISRCTN47552076). Participants were recruited through open invitation with no specific selection criteria were set. Participants were grouped according to their surgical experience. Novices were defined as having no surgical operative experience and less than 4 hours of robotic or laparoscopic simulation training. Intermediate proficiency was set at between 1 and 50 robotic cases independently whilst the expert standard was having performed over 50 robotic cases independently.

Within a simulated operating room environment, participants completed a surgical task and their NTS were assessed. Principle analysis involved the assessment of 59 robotic novices. In addition, intermediate and expert robotic surgeons completed the study. Prior to the study all novice participants were given a 1 hour hands-on-training session on robotic suturing and didactic training on performing a urethrovesical anastomosis. The assessment task comprised a urethrovesical anastomosis using a synthetic dry-lab model (3-Dmed, OH, USA). An "igloo" distributed operating room was used to authentically recreate the OR environment and two actors played the roles of scrub assistant and anaesthetist[14]. Whilst completing the suturing, participants were exposed to four scripted stressor events of increasing magnitude. Firstly, the participants were engaged in direct conversation. Then distracting music was played before the participants were questioned on the state of the patient and potential bleeding. Finally, the patient was simulated to become haemodynamically unstable, which the surgeon was required to manage the scenario whilst completing the procedure. All scenarios were recorded using 3 external rooms cameras in addition to the internal video feed from the robotic endoscopic camera. For the principle analysis, all videos were blindly evaluated post hoc by a panel of expert robotic surgeons and a NTS expert. All participants were assessed by the NTS expert and nested within the expert robotic surgeons with overlap to enable agreement analysis. Secondary analysis of intermediate and expert robotic participants was undertaken by expert robotic surgeons only. All participants were assessed using ICARS and NOTSS.

Statistical Analysis

Categorical data was reported as frequency, n, and percentage. Continuous data is reported as mean + standard deviation (SD).

Given its widespread implementation and validation, NOTSS was defined as the gold standard. Correlation with ICARS was evaluated using a Bland-Altman plot. 95% confidence intervals were corrected for the variance and bias associated with repeated observations[15]. For direct comparison, mean NOTSS and ICARS values were normalised using Z scores. The ability of ICARS to accurately discriminate between participants assessed through the comparison of novice, intermediate and expert participants using the Kruskal Wallis Test. For the purposes of this study we have assumed a direct correlation between surgical experience and technical ability[16]. Global interrater reliability was assessed using Krippendorff's Alpha given the nested rating design. Direct comparison of raters using intraclass correlation coefficient for agreement ($ICC_{agreement}$) coefficients with a two-way random effects model. Internal consistency of the principle domains was tested using Cronbach's alpha. Floors effects were considered present if more than 15% of the novices achieved the minimum score[17]. Ceiling effects could not be assessed given the relatively low numbers of expert participants.

Results

Development of ICARS

During the first phase of development, 45 distinct behaviours constituting robotic NTS were identified and included in the preliminary checklist. The response rate for the Delphi process was 100%. Two rounds of the Delphi process were required to reach a consensus amongst the panel and produce the finalised list of NTS behaviours. The final checklist contained 28 core component behaviours, divided into four domains and seven categories (Figure 1). In addition to generic domains of situation awareness, decision making, task management, leadership and communication and team work, three further domains were felt to warrant inclusion (WHO checklist completion, Console Set Up and Stress and Distractors).

Validation of ICARS Agreement

Feasibility, Acceptability and Educational Impact

All of the expert panel agreed that the use of ICARS for NTS training was important for training in robotic surgery with 86% agreeing that ICARS could be appropriately applied to the robotic surgical environment (Figure 2). Interestingly this was in the context of only 53% having ever used an alternative NTS checklist. 68% of the panel did not think that NTS were currently recognised within robotic surgical training. Only 40% felt an additional guidance sheet would be required prior to use, highlighting its easy-to-use design. There was significant support for the implementation of ICARS both as a learning and assessment tool, 80% agreed that NTS could be actively assessed using ICARS and help identify deficits in NTS. 73% believed ICARS would also promote beneficial discussion and learning follow the case. As a result, 80% of experts agreed that they would be happy to implement ICARS during training.

Validation and Reliability Assessment

73 participants completed the trial composed of 59 novices, 6 intermediate surgeons and 8 expert robotic surgeons. The novices had an average of 30min simulation experience on

open or laparoscopic simulation. None had robotic experience. The intermediate and expert surgeons had mean robotic experience of 5 and 430 cases respectively. Laparoscopic and robotic simulation experience between intermediates and experts were statistically similar. Experts did have greater experience in non-technical skills training although this did not reach significance (Supplementary Table 1).

Comparing ICARS to NOTSS, a high degree of correlation was seen on Bland-Altman plot (Figure 3). The Bland Altman analysis indicates an appropriately narrow 95% confidence interval (z score -0.66 to 0.65) with uniform scatter of plots suggesting good agreement.

The ability to accurately differentiate differing levels of NTS competence was assessed by the comparative analysis of novice (n=59), intermediate (n=6) and expert (n=8) surgeons. The assumption that NTS skill positively correlates with surgical experience was verified through our analysis. Significant differences were found between the 3 groups with experts performing best followed by intermediate participants, p=0.01 (Figure 4).

Krippendorff's Alpha was found to be 0.42 indicating a moderate agreement[18]. However, when compared directly, a greater degree of agreement between raters was seen with a mean ICC of 0.60. Similarly, comparison of ratings from the NTS expert and expert robotic surgeons showed a high degree of agreement ($ICC_{agreement}$ = 0.70).

Likewise, the internal structure of the checklist was found to be reliable (Table 1). All five multi-component categories (Communication & Team Skills; Leadership; Decision Making; Situational Awareness; Stress and Distractors) demonstrated high alpha coefficients (median

= 0.92 range 0.85-0.94) demonstrating that the questions accurately represent the category constructs in each case. On analysis of the individual components, only one failed to support the categories construct (Cronbach alpha <0.70) and would result in an improved alpha for the category if deleted. Following discussion amongst the authors, it was decided to retain it, "Appropriate Interaction with bedside assistant surgeon", given the importance of this behaviour. Absence of any floor effects on analysis of novice participants results further supports content of ICARS.

Discussion

We have successfully developed and provided evidence for the preliminary validation of the first NTS behavioural marker system for robotic surgeons. 28 key behaviours relevant to robotic surgery were identified and organised into a structured behavioural rating system. Expert feedback demonstrated a high degree of support for its acceptability, functionality and educational impact of ICARS. Similarly, the validation study confirmed the reliability of ICARS as an appropriate system for measuring NTS of robotic surgeons. Good interrater reliability was seen despite the large pool of raters. Results correlated closely to the gold standard and were able to distinguish clearly between novice, intermediate and expert participants. Reliability was shown to extend to its internal structure. For all five multi-component categories were found to have very strong relationship with their constituent components. This supports the consistency with which ICARS assesses the various NTS behavioural constructs identified during development.

The use of a simulated scenario to validate ICARS offers a number of benefits. Distributed simulation has been shown to be effective in NTS assessment and enabled homogenous yet full immersive training conditions to be set for all 73 participants[14,19]. Post-hoc video analysis allowed blinded assessment to be undertaken by the panel of robotic and NTS experts. Combined with the high number of participants, we were able to conduct a comprehensive reliability assessment of the ICARS behavioural rating tool. Interestingly this study also highlighted the lack of NTS training currently seen in robotic surgery with almost half the surgeons never having used a NTS checklist. A key development aim of this project was to create a tool that is easily used in the everyday setting. An overall structure similar to establish task specific rating systems was used to maximise familiarity for surgeons[8,9,6]. In addition, more specific detail was given to guide accurate behavioural assessment. Both the feedback from the expert panel and the high degree of agreement found between the experts with and without specialist non-technical skills experience, supporting its usability.

Robotic surgery poses considerable NTS challenges to the surgeon distinct to those of open or even laparoscopic surgery. Face-to-face interaction between the surgeon and his team is greatly reduced and there is an increased reliance on the surgical assistant and scrub team. Furthermore, the surgeon's role changes during the operation as he moves between the bedside and console. Each requires different sets of NTS. These differences were reflected both in the NTS selected during the development of ICARS and in the validation results. ICARS assesses the surgeon both at the bedside and the console. Additionally, ICARS includes a number of unique behaviours not described in generic rating systems including awareness of team members whilst at the console, awareness of the patient status and equipment failure[20,4,3,5]. These NTS idiosyncrasies of robotic surgery result from the both the surgeon's reliance on the robotic system and his detachment from the operating room when sat in the encompassing environment of the console. The unique nature of these skills is further highlighted by comparison of novice, intermediate and expert performances. Experts were substantially better at coordinating the activities from the console, appropriate communication, communicating with the team and managing equipment failure (supplementary figure 1). Interestingly intermediate participants were better at managing the patient at the bedside which may be reflected by their greater experience as bedside assistants.

There were limitations to the study. Use of a simulated OR was appropriate for initial validation however ongoing evaluation of ICARS during live surgery and in other specialities is still required. Similarly, whilst we have demonstrated the reliability of ICARS across the majority of key standards, continuing evaluation is required to measure test stability over time as well as determine appropriate benchmarks for training[21,22]. Assessment using ICARS will provide an objective measure of NTS that can be applied alongside technical skills assessments already include in robotic surgical training[11].

ICARS has been developed as a rating tool to provide objective and structured assessment of NTS during robotic surgery. As the use of robotic surgical systems continue to grow, integration of ICARS within training programmes will provide objective and evidenced-based evaluation of NTS in robotic surgery. The use of ICARS will support structured NTS training and the standardised assessment it provides will enable further research into improving safety and performance in robotic surgery. Currently the only commercially available surgical robot is the Da Vinci System, which this study was based on. The key NTS behaviours identified during this study that comprise ICARS will be equally applicable to the new robotic systems currently in development[23]. As a result, ICARS will be applicable to all robot training, allowing surgeons to directly compare and assess their NTS irrespective of the specific robotic system used.

Conclusions

The first NTS behavioural rating system has been successfully developed for robotic surgery using a structured, evidenced-based approach. Initial validation of the ICARS rating systems demonstrates it to be an effective and reliable tool for assessing NTS during robotic surgery. Ongoing evaluation of ICARS during simulation and real life training will help guide the development of a structured training protocol for NTS in robotic surgery.

Acknowledgments

The authors would like to thank Dr D Murphy, Dr C Callaghan, Dr R Caterwell, Dr S Pandian, Dr R Thurairaja, Dr I Loukopoulos, Dr N Kessaris, Dr J Geh, Dr S Van Rij, Dr K Wong, Dr G Blecher and Dr E Fitzgerald O'Connor for their contributions to the study.

Nicholas Raison acknowledges financial support from the Vattikuti Foundation. Nicholas Raison, Kamran Ahmed and Prokar Dasgupta acknowledge the support of the Medical Research Council Centre for Transplantation and King's College London. Kamran Ahmed and Prokar Dasgupta acknowledge the support of financial support from the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS

Foundation Trust. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health. Prokar Dasgupta also acknowledges the support of the Medical Research Council Centre for Transplantation, London Deanery, London School of Surgery, Guy's and St. Thomas' Charity, European Union, Vattikuti Foundation, and Olympus.

Disclosures

Dr Nicholas Raison reports no conflict of interest. Mr Thomas Wood reports no conflict of interest. Dr Oliver Brunckhorst reports no conflict of interest. Dr Takashige Abe reports no conflict of interest. Miss Talisa Ross reports no conflict of interest. Dr Ben Challacombe reports no conflict of interest. Prof Mohammed Shamim Khan reports no conflict of interest. Dr Giacomo Novara reports no conflict of interest. Dr Nicolo Buffi reports no conflict of interest. Prof Henk Van Der Poel reports no conflict of interest. Dr Craig Mcilhenny reports no conflict of interest. Prof Prokar Dasgupta reports no conflict of interest. Dr Kamran Ahmed reports no conflict of interest.

References:

1. Wood TC, Raison N, Haldar S, Brunckhorst O, McIlhenny C, Dasgupta P, Ahmed K (2016) Training Tools for Nontechnical Skills for Surgeons-A Systematic Review. J Surg Educ. doi:10.1016/j.jsurg.2016.11.017
2. Ahmed K, Jawad M, Dasgupta P, Darzi A, Athanasiou T, Khan MS (2010) Assessment and maintenance of competence in urology. Nat Rev Urol 7 (7):403-413. doi:10.1038/nrurol.2010.81
3. Robertson ER, Hadi M, Morgan LJ, Pickering SP, Collins G, New S, Griffin D, McCulloch P, Catchpole KC (2014) Oxford NOTECHS II: a modified theatre team non-technical skills scoring system. PLoS One 9 (3):e90320. doi:10.1371/journal.pone.0090320

4. Undre S, Sevdalis N, Healey AN, Darzi A, Vincent CA (2007) Observational teamwork assessment for surgery (OTAS): refinement and application in urological surgery. World J Surg 31 (7):1373-1381. doi:10.1007/s00268-007-9053-z
5. Mishra A, Catchpole K, McCulloch P (2009) The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. Qual Saf Health Care 18 (2):104-108. doi:10.1136/qshc.2007.024760
6. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D (2006) Development of a rating system for surgeons' non-technical skills. Med Educ 40 (11):1098-1104. doi:10.1111/j.1365-2929.2006.02610.x
7. Michinov E, Jamet E, Dodeler V, Haegelen C, Jannin P (2014) Assessing neurosurgical non-technical skills: an exploratory study of a new behavioural marker system. J Eval Clin Pract 20 (5):582-588. doi:10.1111/jep.12152
8. Flin R, Patey R (2011) Non-technical skills for anaesthetists: developing and applying ANTS. Best Pract Res Clin Anaesthesiol 25 (2):215-227. doi:10.1016/j.bpa.2011.02.005
9. Mitchell L, Flin R, Yule S, Mitchell J, Coutts K, Youngson G (2013) Development of a behavioural marker system for scrub practitioners' non-technical skills (SPLINTS system). J Eval Clin Pract 19 (2):317-323. doi:10.1111/j.1365-2753.2012.01825.x
10. Hull L, Bicknell C, Patel K, Vyas R, Van Herzeele I, Sevdalis N, Rudarakanchana N (2016) Content Validation and Evaluation of an Endovascular Teamwork Assessment Tool. Eur J Vasc Endovasc Surg 52 (1):11-20. doi:10.1016/j.ejvs.2015.12.044
11. Fisher RA, Dasgupta P, Mottrie A, Volpe A, Khan MS, Challacombe B, Ahmed K (2015) An over-view of robot assisted surgery curricula and the status of their validation. Int J Surg 13:115-123. doi:10.1016/j.ijsu.2014.11.033
12. Challacombe B, Ahmed K, Soomro N, Dasgupta P, Khan MS, Cross W, Weston R, Sanger V, Joyce A, O’Flynn K, Speakman M (2015) British Association Of Urological Surgeons(Baus)Robotic Surgery Curriculum---Guidelines For Training.
13. Mottrie A, Novara G, van der Poel H, Volpe A (2017) ERUS Robotic Curriculum. http://uroweb.org/section/erus/education/. Accessed 09/01/2017
14. Brewin J, Tang J, Dasgupta P, Khan MS, Ahmed K, Bello F, Kneebone R, Jaye P (2015) Full immersion simulation: validation of a distributed simulation environment for technical and non-technical skills training in Urology. BJU Int 116 (1):156-162. doi:10.1111/bju.12875
15. Bland JM, Altman DG (2007) Agreement Between Methods of Measurement with Multiple Observations Per Individual. Journal of Biopharmaceutical Statistics 17 (4):571-582. doi:10.1080/10543400701329422
16. Brunckhorst O, Shahid S, Aydin A, Khan S, McIlhenny C, Brewin J, Sahai A, Bello F, Kneebone R, Shamim Khan M, Dasgupta P, Ahmed K (2015) The Relationship Between Technical And Nontechnical Skills Within A Simulation-Based Ureteroscopy Training Environment. J Surg Educ 72 (5):1039-1044. doi:10.1016/j.jsurg.2015.04.002
17. McHorney CA, Tarlov AR (1995) Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 4 (4):293-307
18. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33 (1):159-174
19. Khan M, Ahmed K, Gavazzi A, Gohil R, Thomas L, Poulsen J, Ahmed M, Jaye P, Dasgupta P (2013) Development and implementation of centralized simulation training: evaluation of feasibility, acceptability and construct validity. BJU Int 111 (3):518-523. doi:10.1111/j.1464-410X.2012.11204.x

20. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D (2006) Development of a rating system for surgeons' non-technical skills. Medical education 40 (11):1098-1104
21. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.) (2014) Standards for educational and psychological testing.
22. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB (2011) Observational tools for assessment of procedural skills: a systematic review. Am J Surg 202 (4):469-480 e466. doi:10.1016/j.amjsurg.2010.10.020
23. Prokar Dasgupta TM, Koon H. Rha, Michael Stark (2017) Robots on the Horizon. Paper presented at the American Urological Association, San Diago, US,

Figure and Table Legends

Figure 1: The Interpersonal and Cognitive Assessment for Robotic Surgery (ICARS) Behavioural Rating System

Figure 2: Mean Expert Agreement with ICARS Category Constructs

Figure 3: Bland Altman Analysis Comparing ICARS with NOTSS

Figure 4: Comparison of Novice, Intermediate and Expert Performance

Table 1: Internal Consistency Analysis for the Principle Categories of ICARS

Supplementary Table 1: Demographic and Surgical Experience of Validation Study Participants