



RAS-NOTECHS: validity and reliability of a tool for measuring non-technical skills in robotic-assisted surgery settings

Julia Schreyer^{1,2} · Amelie Koch¹ · Annika Herlemann³ · Armin Becker³ · Boris Schlenker³ · Ken Catchpole⁴ · Matthias Weigl^{1,5}

Received: 12 October 2020 / Accepted: 26 March 2021
© The Author(s) 2021

Abstract

Background Non-technical skills (NTS) are essential for safe surgical practice as they impact workflow and patient outcomes. Observational tools to measure operating room (OR) teams' NTS have been introduced. However, there are none that account for the specific teamwork challenges introduced by robotic-assisted surgery (RAS). We set out to develop and content-validate a tool to assess multidisciplinary NTS in RAS.

Methodology Stepwise, multi-method procedure. Observations in different surgical departments and a scoping literature review were first used to compile a set of RAS-specific teamwork behaviours. This list was refined and expert validated using a Delphi consensus approach consisting of qualitative interviews and a quantitative survey. Then, RAS-specific behaviours were merged with a well-established assessment tool on OR teamwork (NOTECHS II). Finally, the new tool—RAS-NOTECHS—was applied in standardized observations of real-world procedures to test its reliability (inter-rater agreement via intra-class correlations).

Results Our scoping review revealed 5242 articles, of which 21 were included based on pre-established inclusion criteria. We elicited 16 RAS-specific behaviours from the literature base. These were synthesized with further 18 behavioural markers (obtained from 12 OR-observations) into a list of 26 behavioural markers. This list was reviewed by seven RAS experts and condensed to 15 expert-validated RAS-specific behavioural markers which were then merged into NOTECHS II. For five observations of urologic RAS procedures (duration: 13 h and 41 min), inter-rater agreement for identification of behavioural markers was strong. Agreement of RAS-NOTECHS scores indicated moderate to strong agreement.

Conclusions RAS-NOTECHS is the first observational tool for multidisciplinary NTS in RAS. In preliminary application, it has been shown to be reliable. Since RAS is rapidly increasing and challenges for effective and safe teamwork remain at the forefront of quality and safety of surgical care, RAS-NOTECHS may contribute to training and improvement efforts in technology-facilitated surgeries.

Keywords Robotic-assisted surgery · Surgical teamwork · Human factors · Non-technical skills · Assessment · Da Vinci system

✉ Julia Schreyer
Julia.Schreyer@med.uni-muenchen.de

- ¹ Institute and Clinic for Occupational, Social and Environmental Medicine, University Hospital, LMU Munich, Ziemssenstrasse 1, 80336 Munich, Germany
- ² Institute for Medical Information Processing, Biometry and Epidemiology (IBE), LMU Munich, Munich, Germany
- ³ Department of Urology, University Hospital Grosshadern, Ludwig-Maximilians-University Munich, Munich, Germany
- ⁴ Department of Anesthesia and Perioperative Medicine, Medical University of South Carolina, Charleston, USA
- ⁵ Institute for Patient Safety, University Hospital, Bonn University, Bonn, Germany

In surgery, non-technical skills (NTS) have been shown to be associated with surgeons' technical performance [1–3], quicker crisis resolution [4], rectification of adverse events [5], operative workflow [6], and objective patient outcomes [3]. NTS are defined as 'the cognitive, social and personal resource skills that complement technical skills, and contribute to safe and efficient task performance' (p. 1) [7]. NTS comprise situational awareness, decision-making, leadership, teamwork, and communication [7]. Previous research revealed a relationship between lack of NTS and technical error [2], poor teamwork and operative disruption [8, 9]. Failures in communication are the second most common

contributing factor to surgical incidents [10]. One recent study reported that almost a third of intraoperative incidents were deemed avoidable and associated to failures in NTS [11]. Thus, assessment of NTS and efforts to improve surgical teamwork are essential to improve patient safety in the operating room (OR) [11].

Observational tools are commonly used to quantify and evaluate OR teamwork behaviour and NTS [12]. Several direct observational tools have been introduced to identify and evaluate NTS in ORs [13, 14]. The Oxford Non-Technical Skills (NOTECHS) (II) is a well-established tool that has been applied to various surgical specialities [12]. In their systematic reviews, Li et al. and McMullan et al. concluded that NOTECHS is amongst the observational tools with the highest validity and reliability [12, 13]. However, NOTECHS was designed to measure NTS in conventional surgical teamwork settings and may not be applicable to robotic-assisted surgery (RAS).

In the past decade, there has been a tremendous growth in the use of robotic technologies. Yet, RAS adds unique challenges—both technical and non-technical—compared to conventional surgery [15, 16]. During RAS procedures, one surgeon spends much of the operating time on the console, separated from both the patient and the remaining OR team. This requires different strategies of communication and personal interaction among the OR team [17]. While operating, the console surgeon is not able to see other team members' physical movements and nonverbal answers but has to rely on verbal communication [15]. Other OR team professionals may face challenges with the robot arms obstructing their view of one another. But apart from impeding communication and coordination, RAS also facilitates new modes of nonverbal communication, e.g., with the display as a mediator, providing visual access to all OR team members and allowing for visual communication between the console surgeon and the assistant surgeon [16, 17]. Consequently, available OR teamwork tools may not be sufficiently valid to evaluate NTS in RAS, as they do not take into account RAS-specific behaviours and unique demands [6]. So far, only one observational tool for NTS in RAS has been published: the Interpersonal and Cognitive Assessment for Robotic Surgery (ICARS) [18]. Yet, ICARS merely evaluates NTS of the console surgeon, omitting all other OR professionals' behaviours, i.e., bedside assistant, anaesthetists, and OR nurses [18]. Taking the key role of multi-professional OR teamwork into consideration, there is a strong need for a tool tailored to RAS that comprehensively evaluates the entire multi-professional OR team [2]. Drawing upon a stepwise systematic procedure, we set out to adapt the NOTECHS specifically to RAS procedures with particular consideration to the challenges introduced by robotic technologies.

Aim of the study

Specifically, our study aimed for the following objectives:

- (1) Development of RAS-NOTECHS—an observational tool for the assessment of NTS during RAS that comprises RAS-specific behaviours which are validated by RAS experts, with sub goals of:
 - (1a) Synthesis and compilation of RAS-specific teamwork behaviours through the combination of exploratory observations and a scoping review of the current literature base; with subsequent pre-selection of behavioural markers through internal consensus process;
 - (1b) Refinement and final selection of RAS behavioural markers through expert validation; synthesis of expert validated behavioural markers with finalization of RAS-NOTECHS.
- (2) Establishing initial reliability of RAS-NOTECHS in real-world RAS procedures.

Methods

Design

We drew up a multi-step procedure to develop a tool for measuring NTS in RAS settings and to test this new tool for content validity and reliability. We used a sequential mixed methods design including direct observations, a scoping review, expert interviews, an expert survey, and structured observations.

The study was approved by the Ethics Committee of the Medical Faculty, Munich University (Nr. 19-696). All participants received written and verbal information prior to data collection. Participation was voluntary and written consent was obtained.

Description of NOTECHS II

Our work is based on NOTECHS which was originally used in aviation and was adapted for OR teams by Mishra et al. [14]. It has been previously revised to *Oxford NOTECHS II* [19]. NOTECHS II consists of four behavioural dimensions: leadership and management, teamwork and cooperation, problem-solving and decision-making, and situation awareness. For each, there are three to four categories with a description of the generic skills. Positive and negative

behavioural markers, specific for each OR subteam (surgery, anaesthesia, nursing), are listed for each dimension [14, 19]. A trained observer rates each OR subteam on the four behavioural dimensions. Scores range from 1 ('consistently compromises patient safety and effective teamwork') to 8 ('consistently enhances patient safety and effective teamwork') with 6 being the baseline score ('consistently maintains an effective level of patient safety and teamwork') [19].

Procedure

Our procedure steps are shown in Fig. 1.

Step 1: Exploratory non-standardized observations to identify RAS-specific teamwork behaviours

In order to obtain a preliminary set of behavioural markers and to gain familiarity with the RAS-setting, we first observed 26 live urologic RAS procedures in a university hospital. Both observers (authors: JS, MW) have a behavioural science background (JS in health sciences, MW in psychology, human factors, and teamwork in acute care and surgical settings). Afterwards, 12 procedures (nine urologic, three visceral surgeries) were observed across different hospital sites (one community and three university hospitals). At all sites, da Vinci surgical systems (Intuitive Surgical, Inc., Sunnyvale, CA, USA, models Si and Xi) were used. Observers took unstructured qualitative notes of what they rated as relevant or RAS-specific behaviours and NTS. They based their notes on the observed behaviours, prior knowledge from preliminary research, and on informal conversations with OR team members of the nursing, anaesthetic, and surgical subteams. From the notes, relevant behaviours were identified in a study team internal review process and extracted into an unstructured list of potentially relevant behavioural markers.

Step 2: Scoping literature review on RAS-specific teamwork behaviours

In order to identify RAS-specific behavioural markers, we concurrently conducted a scoping review of the literature. Our aim was to elicit NTS or teamwork behaviours that were described as well as which underlying challenges to NTS are introduced into the OR by surgical robotic systems. Our scoping review procedure was based on the guidelines published by the Joanna Briggs Institute [20].

Inclusion criteria

Articles were selected based on the following pre-established criteria, shown in Table 1.

Search strategy

Articles found in a preliminary search were searched for potential text words (title and abstract), key words, and MeSH-terms. We then searched nine databases (Web of Science, PsycINFO, PubMed, Taylor and Francis, Google Scholar, IEEE Xplore, ACM Digital Library, Science Direct, and Wiley Online Library) for relevant literature. The search strategy consisted of three key concepts: (1) 'operating room' or 'operating room personnel' (i.e., setting or population), (2) the robotic system (i.e., intervention), and (3) 'NTS' (i.e., outcome), the final Pubmed syntax is available (online supplementary file). The search was conducted on April 15 and 16, 2020.

Article selection and data extraction of RAS-specific NTS and challenges

First, all duplicate publications were removed. Two reviewers (JS, MW) then independently performed title/abstract screening and full-text screening. After each step, any disagreements were solved by discussion. We further checked reference lists of included publications

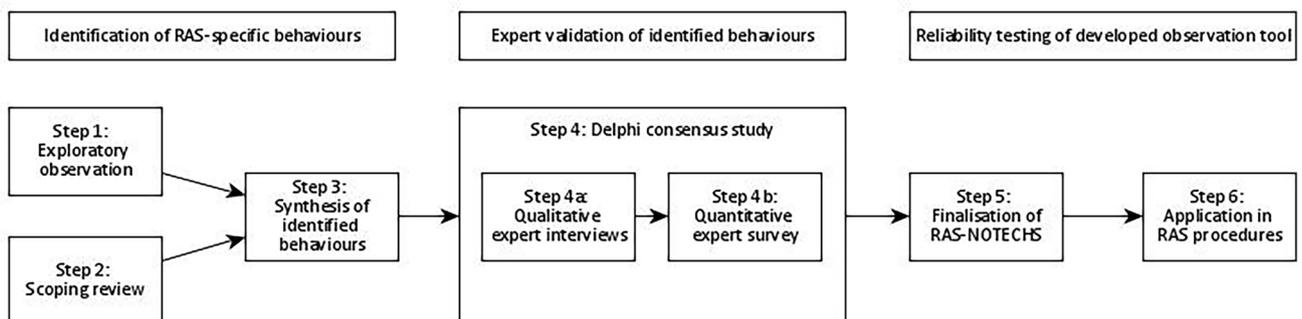


Fig. 1 Study procedure flowchart

Table 1 Inclusion and exclusion criteria for article selection

Inclusion criteria	Exclusion criteria
Topic: non-technical skills or team behaviours in RAS-setting Scientific article of any kind (original study, review, letter, reports) Setting: surgeries facilitated by robotic technology Time frame: Articles published between 1995 and 2020 Language: English or German	Studies on... Technical development of surgical robots or specific parts of surgical robots (technical challenges, designs, ...) Clinical applicability of surgical robots Aetiology/pathology of a disease Therapy options other than surgery (such as medication) Surgical therapy options for specific diseases (options, outcomes, feasibility, clinical trials) Robots other than surgical robots (robotic nurse, therapy) Telemedicine

for additional relevant studies. A data extraction sheet was developed to capture study information, which included first author, year of publication, country of origin, as well as relevant content regarding methodology, surgical specialty, profession, robotic system, RAS-relevant NTS as well as challenges to NTS due to RAS. Data were extracted by the first author (JS).

Step 3: Synthesis of observation results and scoping review

Retrieved behaviours from steps 1 and 2 were combined into a first comprehensive list of behaviours and, where necessary, rephrased. In a study team's internal consensus process (JS, AK, MW, KC), we further assigned all behavioural markers to respective NOTECHS II-dimensions and subteams, according to the definitions provided [14, 21]. We excluded behaviours that were not specific (e.g., 'effective communication is important' [22]).

Step 4: Delphi consensus study with OR experts

In order to further condense the list of behaviour markers and to establish content validity, we applied a Delphi consensus approach [23]. The first round consisted of qualitative interviews and the second of a quantitative survey. Participants were OR personnel (three surgeons, three OR nurses, one anaesthetist) from two surgical academic urology departments of large tertiary care hospitals (about 1000 and 2000 hospital beds, respectively). All participants had broad experience with RAS, ranging from 1.5 to 11 years (surgeons between 8.5 and 11 years, OR nurses between 1.5 and 9 years, and anaesthetist 2.5 years).

Round one—expert interviews

In order to gather expert views on the behaviours found in literature and observation, we conducted six semi-structured expert interviews discussing the list of behavioural markers (retrieved in step 3). In each interview, the list of behavioural markers was evaluated by an expert who was asked to state whether he/she rated listed behaviours as positive or negative with regard to teamwork, and whether it was relevant to RAS, and to provide justification. Afterwards, experts were asked whether they think of any other, not yet listed teamwork behaviours. Transcripts were based on notes taken throughout the interview, and where possible, audio recordings.

Round two—expert survey

Subsequently, interview results (of step 4a) were analysed and the preliminary list of behaviours refined accordingly, i.e., deleting irrelevant items and rephrasing misleading items. This revised list was then utilized for the second round of the Delphi consensus process. The aim of the survey was to find out which behavioural markers were, in the opinion of experts, indicative of good teamwork and would thus be included in RAS-NOTECHS in the next step. Consistent with Hull et al. [24], all experts were asked to rate to which degree each behaviour contributed positively to OR teamwork and patient safety (using a 5-point Likert scale from 1 = 'not at all' to 5 = 'very much'). Both scores for each behavioural marker were combined into a single score (possible range 2–10). Consensus was a priori defined as agreement (combined score 8–10) among > 80% of respondents. In addition to round 1, an additional RAS-urologist took part in round two.

Step 5: Finalisation of RAS-NOTECHS

By adding the behavioural markers left after the consensus study to the NOTECHS II table of behavioural markers, we created RAS-NOTECHS. NOTECHS II's dimensions, descriptions of generic skills, and the original behavioural markers were left unchanged [14].

Step 6: Pilot application in RAS procedures with test for reliability

Finally, RAS-NOTECHS was tested in real-time, standardised observations of RAS procedures in order to assess reliability. In six urologic RAS procedures (one partial nephrectomy, five radical prostatectomies), NTS were rated by two independent observers (JS, MW) simultaneously using the RAS-NOTECHS in order to assess inter-observer reliability. Additionally, a checklist of RAS behavioural markers was filled out in order to find out which of these behaviours were observable. For the observations, the procedures were divided into several phases [*'wheels in to insufflation'*, *'insufflation to surgeon at console'*, *'surgeon at console to surgeon off console'* (further divided into 20 min intervals), *'surgeon off console to closure'*]. RAS-NOTECHS ratings and the behaviour checklist were assessed for each phase, respectively. Intra-class correlation (ICC) estimates with 95% confidence intervals were calculated for the overall RAS-NOTECHS rating as well as for each RAS-NOTECHS dimension (with a single measure, absolute agreement, two-way mixed-effects model). We calculated ICCs for the overall as well as of each RAS-NOTECHS dimension, respectively. Following the definition by Koo, Li [25], ICC values < 0.5, 0.5–0.75, 0.75–0.9, and > 0.9 are indicative of poor, moderate, good, and excellent reliability, respectively. For agreement on individual RAS behaviours, Gwet's AC1 [26] was calculated. Descriptive analyses were conducted to determine how frequently behavioural markers were observed. We used SPSS 25 (IBM Inc., Chicago).

Results

Step 1: Exploratory observation results

After 73 h and 51 min of observer training, we observed a total of 39 h and 51 min of RAS procedures (radical prostatectomies, partial nephrectomies, adrenalectomies, pyeloplasties, sigmoid resections, intestinal resections) using the da Vinci surgical system. 25 observed behaviours were extracted from observation notes and condensed into a preliminary collection of 18 behavioural markers deemed to be relevant to RAS. This preliminary list is available upon request.

Step 2: Scoping review results

Our systematic literature search retrieved 5242 articles. After duplicate removal and abstract screening, 57 articles underwent full-text review. Finally, 21 eligible articles were included for data extraction, elicitation and synthesis of RAS-specific behaviours.

Characteristics of included articles

Table A1 of the online supplementary files shows characteristics of the 21 included articles. All were published between 2004 and 2020 (median: 2017). Eleven studies were conducted in the US, [15, 17, 27–35], six in the UK [18, 36–40], and four in Western/Northern European countries [22, 41–43]. Studies either examined one [18, 22, 28, 32, 41] or several OR professions [15, 17, 27, 29–31, 33–40, 42, 43]. The majority investigated only one surgical specialty [17, 18, 27–30, 33–35, 38]. Eight studies focused on da Vinci surgical systems (Intuitive Inc., CA) [18, 27, 36, 37, 40–43], one examined Brock Rogers Surgical Laprotek System [29].

Challenges to NTS and team behaviours in RAS settings

Results of the scoping review concerning challenges to NTS that are introduced by RAS are shown in Table 2. Sixteen NTS or teamwork behaviours were extracted. Most of these were mentioned in several studies (e.g., 'use of explicit communication' [15, 18, 31, 37, 41, 42]). Fewer of them were reported only in a single publication (e.g., 'surgeon enforces read-back if no confirmation is produced' [41]).

Step 3: Synthesis results

After both preliminary lists from step 1 and step 2 were combined, the collated list of RAS behavioural markers consisted of 26 items. This list, including its respective NOTECHS dimension, the OR subgroup it refers to, as well as behavioural examples from literature or observation, was used to develop an interview guide for the following expert interviews (this preliminary list can be obtained from the authors upon request).

Step 4: Delphi consensus study results

Round one—interview results

All experts deemed the list of behaviours comprehensive. Some behaviours were rephrased, and one behavioural marker was split in two ('console surgeon keeps team engaged by commenting on operative steps, providing educational commentary' was changed to 'console surgeon keeps team engaged by commenting on operative steps')

Table 2 Challenges to NTS during RAS synthesized from scientific literature (including references)

Reason	Challenge
Change in task distribution	Additional tasks for surgeon: responsible for distributing more information to surgical team (e.g., changing instruments) [1] Surgeon can control more instruments; Reduction in task load for bedside team (scrub practitioner, first assistant), this can lead to decreased engagement in procedure and awareness of processes, making the team less responsive to the console surgeon's need for assistance [2–4]
Physical separation of console surgeon and rest of team	Harder for team to hear surgeon's request, communication has to be repeated often (+/ microphone bad) (esp. when immersed) [2, 4–9] Unclear who surgeon is talking to [2] Surgeon cannot see operating table, relies on surgical assistant and scrub team to communicate [2, 3, 10–12] Surgeon cannot see patient, depends on team to inform him, increased coordination and communication demands [2, 7, 12] Surgeon cannot see robot, depends on team to inform him [1, 2] Surgeon is immersed in console, cannot see team (e.g., no visual feedback that message has been received, needs verbal confirmation that request was accepted and fulfilled) [6–8, 11, 12] Surgeon unscrubbed, not at table bedside team has to respond to complication, needs higher shared awareness [2] Immersed in console, tunnel vision, less aware of what others are doing [2, 3, 8, 10, 13, 14] Team cannot see surgeon difficult for team to monitor surgeon's actions and facial cues [2–4, 12, 13] Team members, especially console surgeon, can experience a sense of isolation [7, 8, 13] Harder to interpret directional cues [1, 2, 4, 15, 16] Surgeon and team cannot see each other, cannot communicate via gestures, movements, face to face communication [9, 17–19] Impedes face to face implicit control, Non-verbal communication difficult, reliance on verbal exchanges, changed feedback-loop [1]
Robotic system itself	Robot/additional equipment size space constraints [2, 5, 9, 10] Robotic system: no tactile information [1, 2, 7, 19] Communication via bidirectional device staff-side talks distracting to console surgeon [8]

and 'console surgeon keeps team engaged by conversation', as experts said that unrelated conversations also helped stay engaged). Six behaviours were excluded from the list because the experts either agreed that this behaviour was neither negative nor positive, or that this behaviour was not specific to RAS. After expert interviews, the list comprised 21 behavioural markers. These were then converted into round two's survey.

Round two—survey results

Pre-defined consensus was reached for 15 of the initial 21 survey items. On seven of these, experts agreed 100%.

Step 5: Finalisation of RAS-NOTECHS results

Finally, 15 expert-approved behavioural markers were added to NOTECHS II (p. 6, table of subteam modifiers) [14]. Twelve behavioural markers were assigned to surgeons, five to anaesthesia, and nine to OR nursing (some were also assigned for two or three professions). Four were added to the dimension *leadership and management*, five (three for all subgroups, two additional for nursing subgroup) to *teamwork and cooperation*, one to *problem-solving and*

decision-making, and five to *situation awareness*. The final RAS-NOTECHS including the table of behavioural markers is presented in the online supplement (Tables A2–A4).

Step 6: Descriptive analysis results and inter-rater agreement results

The first pair-wise observation served as a calibration session, after which some small adjustments to RAS-NOTECHS were made. Following this, two trained observers (JS, MW) simultaneously observed 13 h and 41 min of RAS procedures (convenience sample of one partial nephrectomy, four radical prostatectomies; all performed with da Vinci model Xi).

As shown in Fig. 2, some behaviours such as 'uses explicit communication' and 'tries to keep an appropriate level of noise in the OR' were very prevalent (96.2% and 78.9%, respectively). Others such as 'makes sure that a safe distance to the sterile robot is kept' (3.8%), 'proactively positions monitors in coordination with team members' (17.3%) or 'asks who was addressed if uncertain' (1.9%) were rarely observed. The online supplement (Table A5) lists examples for each behaviour.

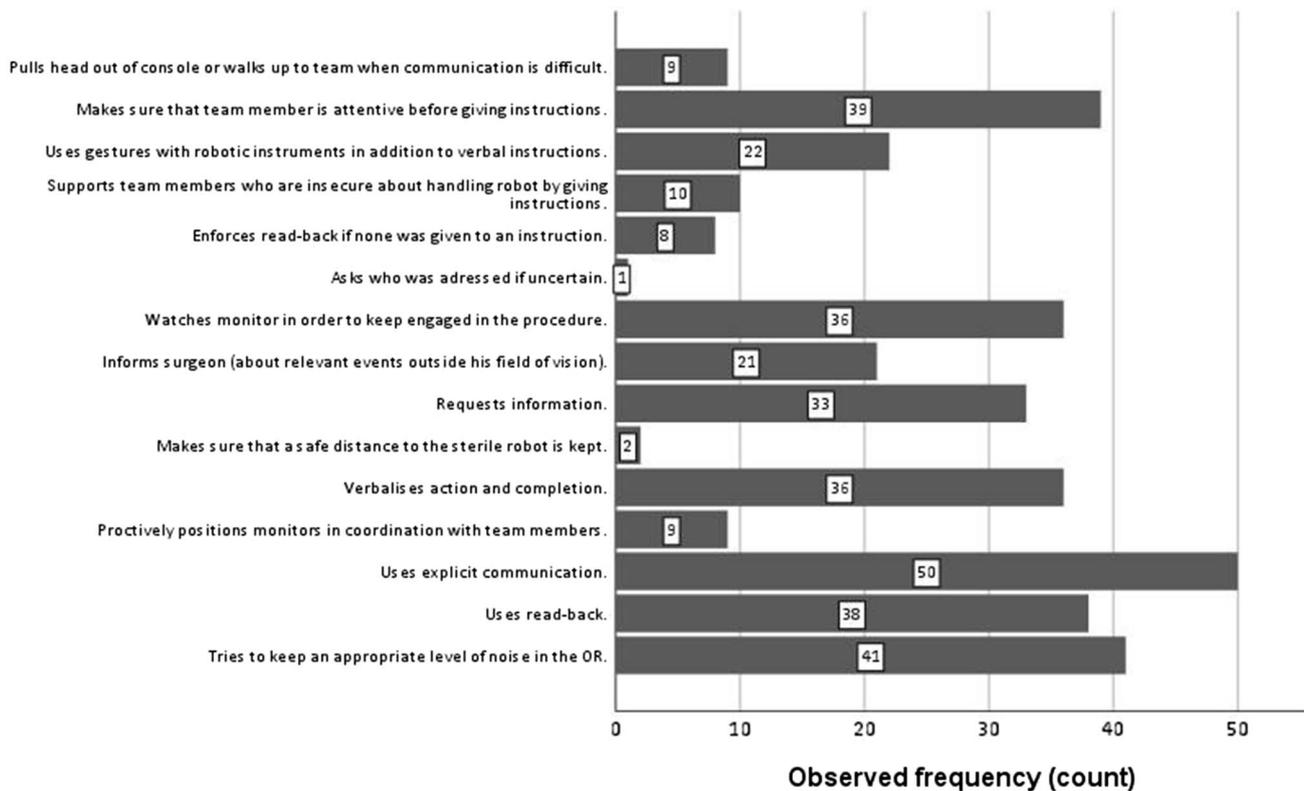


Fig. 2 Observed frequency of RAS-specific behaviours (overall $n = 52$ observation phases)

A moderate degree of correlation was found for the scores of the complete RAS-NOTECHS (ICC 0.687, 95% CI [0.639; 0.729]) as well as for scores of RAS-NOTECHS dimensions *leadership and management* (ICC 0.690, 95% CI [0.583; 0.773]) and *situation awareness* (ICC 0.500, 95% CI [0.360; 0.618]). For the dimensions *teamwork and cooperation* and *problem-solving and decision-making*, we achieved good agreement (ICC 0.812, 95% CI [0.745; 0.863], ICC 0.789, 95% CI [0.715; 0.846], respectively). For inter-rater agreement for the list of individual behaviours we obtained a Gwet's AC1 of 0.831, 95% CI [0.789, 0.874], indicating strong agreement [44, 45].

Discussion

The aim of this study was to develop a tool for measuring teamwork of RAS teams as well as to test its reliability. A systematic, stepwise and mixed-methods procedure was applied. RAS-NOTECHS encompasses relevant behavioural markers for effective OR teamwork in RAS. Thus, our study results contribute to the current literature base in several ways:

First, we introduce a new observational tool that allows assessment of teamwork behaviours in RAS.

Robot-facilitated surgery is rapidly growing and surgical technology is advancing with an increasing demand and use across different procedures [46–48]. Since OR teamwork behaviours are an important factor in safety and quality of delivery of surgical care [2, 3, 5, 11], a reliable and valid tool for assessing NTS in RAS was needed to evaluate multi-professional teamwork in RAS. RAS-NOTECHS is expert validated and has been shown to be a reliable instrument.

Second, drawing upon a scoping review of the current scientific literature base on teamwork in RAS, we systematically identified a comprehensive set of behaviours that are deemed critical for successful collaboration in RAS. Further, this extended set of behaviours was evaluated by clinicians from surgery, nursing, and anaesthesia. The result of the consensus process is a validated set of RAS-critical behaviours. It comprises teamwork behaviours that are assessable, tangible and can be utilized for training, simulation, and teamwork improvement measures [13, 49–52].

Our obtained behavioural markers are to some extent similar to behaviours OR teams show in conventional surgical settings (e.g., open surgery). Yet, we deem that these behaviours are more critical in robotic settings as they may compensate for the unique challenges to NTS and teamwork introduced by RAS [15, 37]. Additionally, most of the identified behavioural markers either applied to the

console surgeon or bedside team alone, or the entire team. Hence, there was no new behavioural marker that exclusively applied to the anaesthesia subteam. This finding is consistent with our scoping review findings: the robot presents challenges mainly for the surgical and nursing subteams by creating physical separation and changing their task load [36–38]. Spatial separation between the anaesthesia subteam and the operating table is also applicable to conventional OR set-ups. However, RAS-specific anaesthesiologic demands remain, e.g., with regard to high relaxation needs or difficult patient positioning, such as Trendelenburg [28].

Despite our focus on potential RAS-related barriers and obstacles to NTS as we elicited in Table 2, it is important to note that there are also relevant teamwork advantages facilitated through RAS. For example, the console surgeon and assistant surgeon can communicate via the screen either using their instruments to point at important areas in situ [15] or by using telestration [53]. Additionally, as each OR team member is able to observe the progression of the surgery on the screen, the team's shared situation awareness increases [16].

While we aimed to identify behaviours indicating good or poor teamwork, we found that some were only applicable if a negative behaviour or event preceded, i.e., actually compensated for obstacles of the robotic setup [15, 35, 37, 40]: for example, 'surgeon pulls head out of console or walks up to operating table if communication with team is difficult' or 'OR team member asks if uncertain who was addressed' can be considered as a necessary rectification of a suboptimal condition. Since we consider OR professional's adequate and effective adaption to evolving challenges an important NTS, these behavioural markers were kept.

We found RAS-NOTECHS to be a reliable teamwork assessment tool throughout the first applications in real-world procedures. Inter-rater agreement was at least moderate in all dimensions which is similar to other studies using NTS observational tools in surgery [19, 49]. Our inter-rater agreement of the individual behavioural markers was strong. The variance of observed behaviours is not surprising: particular behaviours are general and apply to almost each step of a surgery (e.g., 'uses explicit communication'), others are only necessary either in specific phases of a procedure or as compensation for a negative circumstance (e.g., 'proactively positions monitors in coordination with team members', 'asks who was addressed if uncertain'). Future research should confirm reliability and feasibility across different surgical settings, procedures, as well as test for robustness across various observers' backgrounds [54]. In our study, RAS-NOTECHS was content-validated by RAS experts. We did not test for construct nor criterion validity. As these kinds of validity testing are important in order to establish overall validity of RAS-NOTECHS, we recommend

assessing (1) the relationship with other NTS measures and (2) the relationship between RAS-NOTECHS scores and relevant safety and patient outcomes, e.g., number of near-miss events [55].

Third, given the key role of NTS in technology-facilitated surgery and the high pace of technological innovations in surgery, we deem that our multi-step approach may serve as a blueprint for similar attempts to capture teamwork behaviours in high-technology care settings. Given the increasing role of new technology in the OR, multi-disciplinary and inter-professional teamwork will remain crucial for a safe and efficient delivery of surgical care [56, 57]. We propose a systematic procedure for the development of observational tools, combining the current literature base, expert knowledge, and real-life observations.

Limitations

Our study has some limitations. First, we conducted our observations (step 1 and 6) in four university or teaching hospitals in Germany, all working with da Vinci surgical systems. This may limit our proposed behaviours, as possible differences between countries or across robotic systems were not identified. Notwithstanding, our scoping review accounted for variety in RAS practice as the included articles originated from various countries. Just one article specified dealing with a different robotic system, which was, however, similar in setup and resulting challenges [29]. The da Vinci surgical system is currently by far the most widely implemented RAS system worldwide [29], and other surgical systems that are being introduced to the market share its basic setup (e.g., surgeon at console being physically separated) [58, 59]. Therefore, even though RAS-NOTECHS behavioural markers were mainly based on observations and literature on the da Vinci surgical system, they are highly likely to apply to RAS using other surgical systems.

Second, bias inherent to observational studies may have occurred, i.e., participant reactivity and observer bias [60]. Also, since we only observed urologic procedures in one institution for reliability testings (step 6), further tests across specialties and in different institutions are advised.

Third, for expert validation we used a convenience sample of OR staff from only two surgical departments. Our sample size of experts may incur bias concerning institutional practices that are idiosyncratic to specific departments. Our geographically restricted sample of experts and imbalance across professions may have influenced the results of our Delphi consensus study. However, the opinions and behaviours of our experts are in line with the results of our literature review. Future research should investigate how teamwork behaviours in RAS depend upon set-ups (i.e., position of console in relation to anaesthetist and assistant, layout of OR).

Last, we may have possibly missed some relevant literature in our scoping review (i.e., exclusion of grey literature, language restrictions). However, most of the extracted behaviours were consistently mentioned in several of included publications. Additionally, all interviewed expert clinicians deemed the set of behaviours comprehensive. Hence, we assume that all relevant behaviours have been identified. Our scoping review is a first step in examining the literature base of specific NTS in RAS. In the future, methodological quality of included studies may be evaluated and taken into account. In fact, a previous systematic review on a topic comparable to ours concluded that the available literature base on NTS within RAS has methodological limitations [16].

Implications

With regard to implications for future research, RAS-NOTECHS can be applied during observation of real-life procedures and correlated with procedure or patient outcomes in order to expand the current knowledge about the effect of NTS on quality of healthcare delivery. RAS-NOTECHS may assist in assessing effects of teamwork trainings by comparing RAS-NOTECHS scores before and after a training session (e.g., through educational interventions, simulations [61]). In healthcare settings other than RAS, NTS training has been shown to improve NTS performance during simulation or in real-life care [52, 62, 63] and was associated with improvements in patient outcomes [63].

Future research should consider potential influences on RAS teamwork that we did not specifically address in the development and validation of RAS-NOTECHS, such as low- vs. high-complexity procedures, team familiarity (i.e., prior experience of working together, fixed RAS teams), and dealing with intraoperative events or disruptions [64].

With regard to implications for surgical practice, the set of RAS-specific behaviours itself can serve as a basis for teaching teamwork in RAS. We found that OR staff already perform most of the identified behaviours in open surgery, but to a lesser extent compared to RAS, implicating that almost no new skills need to be taught in RAS-NTS training. In fact, our findings advocate that existent skills need to be considered and targeted in RAS-training and education. To this end, already existing and applied training approaches might be adapted [62, 65, 66].

Our new knowledge might be taken into account for the future design of robotic surgical technologies in order to facilitate intraoperative teamwork. We found that many RAS-specific behaviours are necessary to compensate for challenges introduced by the surgical robot, such as obstructed views between console surgeon and the rest of team, bulky equipment [35], difficulties of verbal communication, or low situation awareness [37]. By applying

human-centred design concepts, developers should consider how they can prevent behaviours that compensate for suboptimal conditions and how to design robotic technologies that foster smooth collaboration within the OR team.

Conclusions

In our study, we introduce RAS-NOTECHS—the first behavioural rating system for multidisciplinary NTS in RAS. We deployed a stepwise mixed-methods approach combining the current evidence base, expert knowledge, and real-life observations. RAS-NOTECHS is reliable and can be used to identify RAS-specific behavioural markers in real-life procedures. Our study provides assessment methodologies for future research investigating the role of NTS for safety and quality of delivery of surgical care in technology-facilitated teamwork.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00464-021-08474-2>.

Acknowledgements This work was part of the Master Thesis Requirements of Julia Schreyer (LMU University, Munich). We thank all surgeons, nurses, and anaesthetists who participated in our study and allowed us to observe, let us interview them, and/or filled out our survey. Dr Catchpole is supported under grant number HS026491-01 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services (HHS). The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of AHRQ. Readers should not interpret any statement in this report as an official position of AHRQ or of HHS.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was partly funded by the Munich Centre for Health (MC-Health, Grant Recipients Matthias Weigl and Amelie Koch).

Declarations

Disclosure M.Sc. Julia Schreyer, M.Sc. Amelie Koch, Prof. Dr. Annika Herlemann, Prof. Dr. Armin Becker, Prof. Dr. Boris Schlenker, Dr. Ken Catchpole, and Dr. Matthias Weigl have no conflict of interest or financial ties to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Pfandler M, Stefan P, Mehren C, Lazarovici M, Weigl M (2019) Technical and nontechnical skills in surgery: a simulated operating room environment study. *Spine (Phila Pa 1976)* 44(23):E1396–E400. <https://doi.org/10.1097/brs.0000000000003154>
2. Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N (2012) The impact of nontechnical skills on technical performance in surgery: a systematic review. *J Am Coll Surg* 214(2):214–230. <https://doi.org/10.1016/j.jamcollsurg.2011.10.016>
3. Schmutz JB, Meier LL, Manser T (2019) How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: a systematic review and meta-analysis. *BMJ Open* 9(9):e028280. <https://doi.org/10.1136/bmjopen-2018-028280>
4. Doumouras AG, Hamidi M, Lung K, Tarola CL, Tsao MW, Scott JW et al (2017) Non-technical skills of surgeons and anaesthetists in simulated operating theatre crises. *Br J Surg* 104(8):1028–1036. <https://doi.org/10.1002/bjs.10526>
5. Fecso AB, Kuzulugil SS, Babaoglu C, Bener AB, Grantcharov TP (2018) Relationship between intraoperative non-technical performance and technical events in bariatric surgery. *Br J Surg* 105(8):1044–1050. <https://doi.org/10.1002/bjs.10811>
6. Gjeraa K, Spanager L, Konge L, Petersen RH, Østergaard D (2016) Non-technical skills in minimally invasive surgery teams: a systematic review. *Surg Endosc* 30(12):5185–5199. <https://doi.org/10.1007/s00464-016-4890-1>
7. Flin R, O'Connor P, Crichton M (2008) Safety at the sharp end: a guide to non-technical Skills. Ashgate, Aldershot
8. Gillespie BM, Harbeck E, Kang E, Steel C, Fairweather N, Chaboyer W (2017) Correlates of non-technical skills in surgery: a prospective study. *BMJ Open* 7(1):e014480. <https://doi.org/10.1136/bmjopen-2016-014480>
9. Al-Hakim L (2011) The impact of preventable disruption on the operative time for minimally invasive surgery. *Surg Endosc* 25(10):3385–3392. <https://doi.org/10.1007/s00464-011-1735-9>
10. Mushtaq F, O'Driscoll C, Smith F, Wilkins D, Kapur N, Lawton R (2018) Contributory factors in surgical incidents as delineated by a confidential reporting system. *Ann R Coll Surg Engl* 100(5):401–405. <https://doi.org/10.1308/rcsann.2018.0025>
11. Siu J, Maran N, Paterson-Brown S (2016) Observation of behavioural markers of non-technical skills in the operating room and their relationship to intra-operative incidents. *Surgeon* 14(3):119–128. <https://doi.org/10.1016/j.surge.2014.06.005>
12. Li N, Marshall D, Sykes M, McCulloch P, Shalhoub J, Maruthappu M (2018) Systematic review of methods for quantifying teamwork in the operating theatre. *BJS Open* 2(2):42–51. <https://doi.org/10.1002/bjs.5.40>
13. McMullan RD, Urwin R, Sunderland N, Westbrook J (2020) Observational tools that quantify nontechnical skills in the operating room: a systematic review. *J Surg Res* 247:306–322. <https://doi.org/10.1016/j.jss.2019.10.012>
14. 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. <https://doi.org/10.1136/qshc.2007.024760>
15. Catchpole K, Bisantz A, Hallbeck MS, Weigl M, Randell R, Kosack M et al (2019) Human factors in robotic assisted surgery: lessons from studies “in the Wild.” *Appl Ergon* 78:270–276. <https://doi.org/10.1016/j.apergo.2018.02.011>
16. Gillespie BM, Gillespie J, Boorman RJ, Granqvist K, Stranne J, Erichsen-Andersson A (2020) The impact of robotic-assisted surgery on team performance: a systematic mixed studies review. *Hum Factors*. <https://doi.org/10.1177/0018720820928624>
17. Tiferes J, Hussein AA, Bisantz A, Kozlowski JD, Sharif MA, Winder NM et al (2016) The loud surgeon behind the console: understanding team activities during robot-assisted surgery. *J Surg Educ* 73(3):504–512. <https://doi.org/10.1016/j.jsurg.2015.12.009>
18. Raison N, Wood T, Brunckhorst O, Abe T, Ross T, Challacombe B et al (2017) Development and validation of a tool for non-technical skills evaluation in robotic surgery—the ICARS system. *Surg Endosc* 31(12):5403–5410. <https://doi.org/10.1007/s00464-017-5622-x>
19. Robertson ER, Hadi M, Morgan LJ, Pickering SP, Collins G, New S et al (2014) Oxford NOTECHS II: a modified theatre team non-technical skills scoring system. *PLoS ONE* 9(3):e90320. <https://doi.org/10.1371/journal.pone.0090320>
20. Peters MDJ GC, McInerney P, Munn Z, Tricco AC, Khalil, H. Chapter 11: Scoping reviews (2020 version). In: Aromataris E, Munn Z (eds) *JBI manual for evidence synthesis*. JBI. 2020. <http://doi.org/https://doi.org/10.46658/JBIMES-20-12>. <https://reviewersmanual.joannabriggs.org/>
21. Flin R, Martin L, Goeters KM, Hoermann H, Amalberti R, Valot C et al (2003) Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Hum Factors Aerosp Saf* 3:97–119
22. Myklebust MV, Storheim H, Hartvik M, Dysvik E (2020) Anesthesia professionals' perspectives of teamwork during robotic-assisted surgery. *Aorn J* 111(1):87–96. <https://doi.org/10.1002/aorn.12897>
23. Iqbal S, Papon-Young L (2009) The Delphi method. A step-by-step guide. *Psychologist* 22:598–601
24. Hull L, Arora S, Kassab E, Kneebone R, Sevdalis N (2011) Observational teamwork assessment for surgery: content validation and tool refinement. *J Am Coll Surg* 212(2):234–43.e1–5. <https://doi.org/10.1016/j.jamcollsurg.2010.11.001>
25. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15(2):155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
26. Gwet K (2002) Kappa statistic is not satisfactory for assessing the extent of agreement between raters. *Stat Methods Inter-rater Reliabil Assess* 1:1–6
27. Pelikan HR (2018) “What’s going on there?” Negotiating common ground in robotic vs. open surgery: a comparison of surgeon-initiated requests for action in open and robotic surgery. University of Twente, Enschede
28. Hsu RL, Kaye AD, Urman RD (2013) Anesthetic challenges in robotic-assisted urologic surgery. *Rev Urol* 15(4):178–184. <https://doi.org/10.3909/riu0589>
29. Cao CG, Taylor H (2004) Effects of new technology on the operating room team. Department of Mechanical Engineering, Tufts University, Medford, pp 309–12
30. Higuchi T, Gettman M (2011) Robotic instrumentation, personnel and operating room setup. In: Su LM (ed) *Atlas of robotic urologic surgery*. Humana Press, New York, pp 15–30. https://doi.org/10.1007/978-1-60761-026-7_2
31. Lai F, Entin E (2005) Robotic surgery and the operating room team. *Proc Hum Factors Ergon Soc Annu Meet* 49(11):1070–1073. <https://doi.org/10.1177/154193120504901115>
32. Wastler KE (2015) Robotic surgical and anesthesia communication tool. *J Robot Surg* 9(1):97–98. <https://doi.org/10.1007/s11701-014-0494-8>
33. Payne TN, Pitter MC (2011) Robotic-assisted surgery for the community gynecologist: can it be adopted? *Clin Obstet Gynecol* 54(3):391–411. <https://doi.org/10.1097/GRF.0b013e31822b4998>
34. Satchidanand A, Higginbotham J, Bisantz A, Aldhaam NA, Elsayed AS, Kannappan R et al (2019) Vague swirling ends in no communication: referencing strategies during robot assisted surgery. *J Urol* 201(4):E288–E289

35. Raheem S, Ahmed YE, Hussein AA, Johnson A, Cavuoto L, May P et al (2018) Variability and interpretation of communication taxonomy during robot-assisted surgery: do we all speak the same language? *BJU Int* 122(1):99–105. <https://doi.org/10.1111/bju.14150>
36. Randell R, Greenhalgh J, Hindmarsh J, Honey S, Pearman A, Alvarado N et al (2019) How do team experience and relationships shape new divisions of labour in robot-assisted surgery? A realist investigation. *Health*. <https://doi.org/10.1177/1363459319874115>
37. Randell R, Honey S, Hindmarsh J, Alvarado N, Greenhalgh J, Pearman A et al (2017) A realist process evaluation of robot-assisted surgery: integration into routine practice and impacts on communication, collaboration and decision-making. *Health Serv Deliv Res*. <https://doi.org/10.3310/hsdr05200>
38. Alvarado N, Honey S, Greenhalgh J, Pearman A, Dowding D, Cope A et al (2017) Eliciting context-mechanism-outcome configurations: experiences from a realist evaluation investigating the impact of robotic surgery on teamwork in the operating theatre. *Evaluation* 23(4):444–462. <https://doi.org/10.1177/1356389017732172>
39. Gill A, Randell R (2016) Robotic surgery and its impact on teamwork in the operating theatre. *J Perioper Pract* 26(3):42–45. <https://doi.org/10.1177/175045891602600303>
40. El-Hamamsy D, Walton TJ, Griffiths TRL, Anderson ES, Tincello DG (2020) Surgeon-team separation in robotic theaters: a qualitative observational and interview study. *Female Pelvic Med Reconstr Surg* 26(2):86–91. <https://doi.org/10.1097/spv.0000000000000829>
41. Nyssen A-S, Blavier A (2009) Verbal communication as a sign of adaptation in socio-technical systems: the case of robotic surgery. In: 9th Bi-annual international conference on naturalistic decision making (NDM9), vol 9. pp 267–72. <http://doi.org/https://doi.org/10.14236/ewic/NDM2009.39>
42. Almeras C (2019) Operating room communication in robotic surgery: place, modalities and evolution of a safe system of interaction. *J Visc Surg* 156(5):397–403. <https://doi.org/10.1016/j.jvisurg.2019.02.004>
43. Sergeeva A, Huysman M, Faraj S (2015) Transforming work practices of operating room teams: the case of the Da Vinci robot. In: Transforming surgery practice with the Da Vinci robot thirty sixth international conference on information systems. Fort Worth.
44. Viswanathan M, Berkman N (2011) Development of the RTI Item Bank on risk of bias and precision of observational studies [Internet] Appendix A, AC1 Statistic. Agency for Healthcare Research and Quality (US), Rockville. <https://www.ncbi.nlm.nih.gov/books/NBK82266/>.
45. Fleiss J (1971) Measuring nominal scale agreement among many raters. *Psychol Bull* 76:378–382. <https://doi.org/10.1037/h0031619>
46. Peters BS, Armijo PR, Krause C, Choudhury SA, Oleynikov D (2018) Review of emerging surgical robotic technology. *Surg Endosc* 32(4):1636–1655. <https://doi.org/10.1007/s00464-018-6079-2>
47. Perez RE, Schwaitzberg SD (2019) Robotic surgery: finding value in 2019 and beyond. *Ann Laparosc Endosc Surg*. <https://doi.org/10.21037/ales.2019.05.02>
48. Sheetz KH, Claflin J, Dimick JB (2020) Trends in the adoption of robotic surgery for common surgical procedures. *JAMA Netw Open* 3(1):e1918911. <https://doi.org/10.1001/jamanetworkopen.2019.18911>
49. 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. <https://doi.org/10.1007/s00268-007-9053-z>
50. Yule S, Flin R, Maran N, Rowley D, Youngson G, Paterson-Brown S (2008) Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World J Surg* 32(4):548–556. <https://doi.org/10.1007/s00268-007-9320-z>
51. Robertson JM, Dias RD, Yule S, Smink DS (2017) Operating room team training with simulation: a systematic review. *J Laparoendosc Adv Surg Tech A* 27(5):475–480. <https://doi.org/10.1089/lap.2017.0043>
52. Yule S, Parker SH, Wilkinson J, McKinley A, MacDonald J, Neill A et al (2015) Coaching non-technical skills improves surgical residents' performance in a simulated operating room. *J Surg Educ* 72(6):1124–1130. <https://doi.org/10.1016/j.jsurg.2015.06.012>
53. Ali MR, Loggins JP, Fuller WD, Miller BE, Hasser CJ, Yellowlees P et al (2008) 3-D telestration: a teaching tool for robotic surgery. *J Laparoendosc Adv Surg Tech A* 18(1):107–112. <https://doi.org/10.1089/lap.2007.0051>
54. Russ S, Hull L, Rout S, Vincent C, Darzi A, Sevdalis N (2012) Observational teamwork assessment for surgery: feasibility of clinical and nonclinical assessor calibration with short-term training. *Ann Surg* 255(4):804–809. <https://doi.org/10.1097/SLA.0b013e31824a9a02>
55. Manuguerra A, Mazeaud C, Hubert N, Eschwège P, Roumiguié M, Salleron J et al (2020) Non-technical skills in robotic surgery and impact on near-miss events: a multi-center study. *Surg Endosc*. <https://doi.org/10.1007/s00464-020-07988-5>
56. Manser T (2009) Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiol Scand* 53(2):143–151. <https://doi.org/10.1111/j.1399-6576.2008.01717.x>
57. Leonard M, Graham S, Bonacum D (2004) The human factor: the critical importance of effective teamwork and communication in providing safe care. *Qual Saf Health Care* 13(Suppl 1):i85-90. https://doi.org/10.1136/qhc.13.suppl_1.i85
58. Rao PP (2018) Robotic surgery: new robots and finally some real competition! *World J Urol* 36(4):537–41. <https://doi.org/10.1007/s00345-018-2213-y>
59. Brodie A, Vasdev N (2018) The future of robotic surgery. *Ann R Coll Surg Engl* 100(Suppl 7):4–13. <https://doi.org/10.1308/rcsann.supp2.4>
60. Paradis E, Sutkin G (2017) Beyond a good story: from Hawthorne effect to reactivity in health professions education research. *Med Educ* 51(1):31–39. <https://doi.org/10.1111/medu.13122>
61. Lynch A (2020) Simulation-based acquisition of non-technical skills to improve patient safety. *Semin Pediatr Surg* 29(2):150906. <https://doi.org/10.1016/j.sempedsurg.2020.150906>
62. Nicksa GA, Anderson C, Fidler R, Stewart L (2015) Innovative approach using interprofessional simulation to educate surgical residents in technical and nontechnical skills in high-risk clinical scenarios. *JAMA Surg* 150(3):201–207. <https://doi.org/10.1001/jamasurg.2014.2235>
63. Weaver S, Dy S, Rosen M (2014) Team-training in healthcare: a narrative synthesis of the literature. *BMJ Qual Saf* 23(5):359–372. <https://doi.org/10.1136/bmjqs-2013-001848>
64. Koch A, Burns J, Catchpole K, Weigl M (2020) Associations of workflow disruptions in the operating room with surgical outcomes: a systematic review and narrative synthesis. *BMJ Qual Saf*. <https://doi.org/10.1136/bmjqs-2019-010639>
65. Catchpole KR, Dale TJ, Hirst DG, Smith JP, Giddings TA (2010) A multicenter trial of aviation-style training for surgical teams. *J Patient Saf* 6(3):180–186. <https://doi.org/10.1097/PTS.0b013e3181f100ea>
66. Stewart-Parker E, Galloway R, Vig S (2017) S-TEAMS: a truly multiprofessional course focusing on nontechnical skills to improve patient safety in the operating theater. *J Surg Educ* 74(1):137–144. <https://doi.org/10.1016/j.jsurg.2016.06.020>