Effects of a Brief Team Training Program on Surgical Teams' Nontechnical Skills: An Interrupted Time-Series Study

Author

Gillespie, Brigid M, Harbeck, Emma, Kang, Evelyn, Steel, Catherine, Fairweather, Nicole, Panuwatwanich, Kriengsak, Chaboyer, Wendy

Published

2017

Journal Title

Journal of Patient Safety

Version

Accepted Manuscript (AM)

DOI

10.1097/PTS.0000000000000361

Effect of a brief team training program on surgical teams' non-technical skills: An interrupted time-series study

Authors:

[1,2,3]*Brigid M. Gillespie PhD, RN, FACORN; [4]Emma Harbeck BS (Hons), [3]Evelyn Kang MHP (Hons), RN, [5]Catherine Steel MN, RN; [5]Nicole Fairweather FANZCA, MBBS; [6]Kriengsak Panuwatwanich PhD, & [1]Wendy Chaboyer PhD RN

[1]School of Nursing & Midwifery, Griffith University, Gold Coast, QLD, AUSTRALIA

[2]Gold Coast University Hospital and Health Service, QLD, AUSTRALIA

[3]National Centre of Research Excellence in Nursing, Menzies Health Institute of Queensland, Griffith University, Gold Coast, QLD, AUSTRALIA

[4]School of Applied Psychology, Griffith University, Gold Coast Campus, QLD, AUSTRALIA

[5]Division of Surgery, Princess Alexandra Hospital, Woolloongabba, Brisbane, QLD, AUSTRALIA

[6] Griffith School of Engineering Griffith University, Parklands Drive, Gold Coast Campus QLD, AUSTRALIA

***Corresponding author:**

Professor Brigid Gillespie: b.gillespie@griffith.edu.au

Funding Statement

Brigid M. Gillespie acknowledges the financial support of the Australian Research Council, Early Career Discovery Fellowship Scheme and the NHMRC Centre for Excellence in Nursing Research.

Authors' Contributions

BMG conceived of the study, assisted in data analysis, interpreted results and drafted the manuscript. EH performed data analysis and assisted in interpretation. WC and KS assisted in analysis and interpretation. EK, CS, and NF assisted in recruitment and interpretation. All authors participated in the coordination of the study and read and approved the final manuscript.

Effect of a brief team training program on surgical teams' non-technical skills: An interrupted time-series study

Abstract

Background: Up to 60% of adverse events in surgery are the result of poor communication and teamwork. Non-technical skills in surgery (NOTSS) are critical to the success of surgery and patient safety. The study aim was to evaluate the effect of a brief team training intervention on teams' observed NOTSS.

Methods: Pretest-posttest interrupted time series design with statistical process control analysis was used to detect longitudinal changes in teams' NOTSS. We evaluated NOTSS using the revised NOTECHS weekly over 20-25 weeks before and after implementation of a team training program.

Results: We observed 179 surgical procedures with cardiac, vascular, upper gastro-intestinal, and hepatobiliary teams. Mean posttest NOTECHS scores increased across teams, showing special cause variation. There were also significant before-and-after improvements in NOTECHS scores in respect to professional role, and in the use of the Surgical Safety Checklist.

Conclusions: Our results suggest associated improvements in teams' NOTSS following implementation of the team training program.

Words: 154

Key words: teamwork; intervention; patient safety; operating room; statistical process control.

Introduction

Effective teamwork in surgery is recognised as a critical determinant of safe and efficient performance. Surgery is a complex environment; up to 60% of all adverse events occur in surgery, with 33% resulting in permanent disability, and up to 13% resulting in death [1-3]. As a high risk area, surgery may be more vulnerable to failures in communication and teamwork. Investigating the social and cognitive skills that describe team cohesion have altered the way in which behavioural markers of surgical team performance are understood and measured. These skills, termed non-technical skills in surgery (NOTSS), capture the interpersonal (i.e., communication, teamwork and leadership), and cognitive skills (i.e., decision making and situational awareness); essential safe clinical practice [4]. Evaluation of team NOTSS has become an important focus in patient safety research because poor teamwork and communication errors, and cultural and hierarchical barriers contribute to safety failures [5, 6]. Emerging evidence also suggests that NOTSS impact positively on; a) technical performance [7, 8], b) process improvements, e.g., on-time operating list starts [7], c) antibiotic and venous thromboembolism prophylaxis [8]; and, d) reductions in surgical morbidity and mortality [8, 9]. Also critique of NOTSS facilitates improvement by encouraging both self-reflection and team reflection, identifying training needs, and informing team training approaches [10]. In this study, our aim was to implement a brief team training program based on situational awareness principles that could be readily used in practice.

Methods

We used a pretest-posttest interrupted time series design to assess process changes in NOTSS associated with a short team training intervention. The pre-and-post observation periods lasted 20-25 weeks for each phase. This design was 'pragmatic' because it measured the benefits of situation awareness in observed routine clinical practice [11].

Setting and participants

The setting was a 780-bed quaternary centre in Queensland Australia specializing in all surgeries except for paediatrics and obstetrics/gynaecology. Within the 21 operating room (OR) complex, over 21,000 surgeries are performed each year. Four surgical teams were purposively selected by speciality, i.e., Vascular, Upper Gastro-Intestinal (GI), Hepatobiliary (HPB) and Cardiac. Selection of teams was based on staffs' willingness to participate in the study (including the team training program). Surgical procedures were deliberately chosen to ensure variety in case complexity and types of procedures within specialties. Most surgeries observed were elective. Teams were assigned based on the speciality consultant surgeon and team members were evaluated weekly for 20-25 time weeks for pre-and post-intervention phases (i.e., 40-50 weeks in total). Teams consisted of the following core members; Surgical Consultant or Registrar, Anaesthetic Consultant or Registrar, and Instrument and Circulating Nurses. Core membership was defined as having worked together, weekly or fortnightly, for a minimum of three months.[12] Among teams, there was no overlap of surgeons and their registrars; however there was some overlap with the nursing staff, many of whom, worked across 2-3 surgical specialties. The same team members were observed during the pretest and posttest phases.

Intervention

The development of the team training program, coined '**TEAM**ANATOMY' was informed by crew resource management principles and senior clinician input. The program was delivered using a one hour DVD. The program's theoretically underpinnings focused on the individual's situational awareness (as described by Endsley [13], and shared situational awareness as described by Wright et al. [14]. 'Situational awareness' describes an individual's mindfulness and understanding of 'big picture' information in an ever-changing environment. Endsley [13] proposed three distinct levels of situational awareness based on perception, comprehension, and understanding. While situational awareness is critical to the performance of individuals, it is also essential for the team as a collective. 'Shared situational awareness' refers to an appreciation team members have for others' tasks and roles, and share common perspectives of goals and events [14].

In developing **TEAM**ANATOMY, specific attention was given to individual and team factors that inhibit individual and shared situational awareness in two filmed clinical scenarios. The DVD consisted of three parts; 1) individual and shared situational awareness theory, 2) filmed simulation preoperative patient sign in, and 3) filmed simulation of time out procedure. Components of the program are described in **Supplementary file 1.** The training DVD was distributed to core team members of surgical and anaesthetic medical and nursing staff that regularly worked in the four teams.

Measures and scoring

We assessed the effect of the team training intervention on work processes using the Non-technical skill assessment (Revised NOTECHS) [15]. The 23 items on the revised NOn-TECHnical Skills (NOTECHS) [15] tool examines behaviour referencing five factors; a) Communication and interaction, b) Vigilance/situation awareness, c) Team skills, d) Leadership and management skills, and e) Decision making. Descriptors for these domains are offered elsewhere [15]. A separate NOTECHS scale is provided for each surgical role; Anaesthetist, Surgeon and Instrument Nurse. The NOTECHS contains a 7-point rating scale for each factor, (*1-not done, 2- done poorly, 3-done, 4-done adequately, 5-done well and 6-done very well, 0-not applicable*) [15]. The "not applicable" option meant that a specific item was not relevant or could not be rated on the basis that the behaviour was not observed during the observation period. We replaced 'not applicable' scores with the group mean scores and inverted these scores into an individual item mean. Scores of 1–6 for each item were averaged using the total number of items for that factor. Since all factors had equal importance across each surgical role, after summing each factor we standardized the total scores using the number of items (23) in the scale creating an individual NOTECHS total score (range 1-23). The total team score was calculated using the mean of all individual team members NOTECHS total scores (range 1-23). Higher scores indicated better overall performance. We also calculated the mean NOTECHS scores based on professional role (i.e., surgeon, anaesthetist, nurse) for the pretest and posttest periods. Each observation time point included 2-3 combined observations based on the average scores for each individual team member role (Surgeon Consultant or Registrar, Anaesthetic Consultant or Registrar, and Scrub/Scout Nurse(s)). Reliability and validity of the NOTECHS tool was confirmed by Sevdalis et al.[15]. To ensure methodological rigour in observations our current study, inter-rater checks of 10% of cases were performed by the lead author.

As part of the team training program focused on completion of the Surgical Safety Checklist (SSC), use of the SSC was observed. We collected information about the extent to which an attempt was made to complete the three phases of the checklist (i.e., sign-in, timeout and sign-out) (yes/no). Field notes were made in relation to item completion and participation in the checks.

Data collection

Pretest-posttest observational data for both phases occurred over 40-50 weeks during 2013-15. A trained observer who was an experienced OR nurse received specific training in observational research and use of the revised NOTECHS tool. Observations started when the patient entered the OR (prior to anaesthesia) and ended when the patient left the OR. During each surgical procedure, the observer was positioned away from the operating table, with each team member and all doors in view, ensuring a prime viewing position. For each procedure, the observer documented field notes to describe contextual issues, e.g., clinical events such as bleeding. A second observer was present for 14 procedures to test the observational data collection tool, and to establish interrater reliability which ranged from 86%-95% across the two phases.

Ethics approvals were given by the hospital and the university. Participants signed a consent form and were advised of their right to confidentiality and anonymity, and to withdraw at any time. Consent from staff was renegotiated throughout the observational period. Patients whose operations were observed were informed of the possibility of observations taking place and given the opportunity to decline if they wished.

Data Analysis

Data were entered into IBM SPSS Statistics (V 22, New York, NY, USA) and cleaned, and a random sample of 20% was checked for accuracy. Missing data analyses (using SPSS) for the 24 items ranged from 0.6-1.1%, therefore no missing data transformations were performed. Descriptive statistics included number of observations by professional role, operative time, patient ASA[1], and checklist use. One-way analysis of variance was used to compare pretest differences in total NOTECHS scores among surgeons, anaesthetists and nurses. Paired *t*-tests were used to compare before-and-after differences in total NOTECHS scores relative to professional role. The Chi squared (χ^2) test was used to compare before-and-after observed and expected frequencies in relation to checklist compliance. For these analyses, 95% confidence intervals (CI) were used and a *p*-value <.05 was significant

We used statistical process control (SPC) [11, 12] to analyse team NOTECHS scores by surgical speciality. SPC identifies the variability inherent in any process to enable clinicians and researchers to distinguish small, random fluctuations that are normal, (i.e., common cause variation) from unnatural variations stemming from a specific source, (i.e., special cause variation) [16]. When there is no special cause variation, the process is in control, (i.e., stable). Bennyan et al. [16] advises that 20-25 data points are needed to be confident about distinguishing between special cause and common cause variation. Fewer

[1] ASA: American Society of Anaesthesiologists, measure of patient acuity.

than 20 data points can lead to a high risk of a Type II error [17]. The control charts generated were based on continuous data: therefore both the X-bar (for mean) chart and S-bar (for standard deviation [SD]) charts were used to detect special cause variation [17]. The horizontal centre line of the control chart represents the mean of all subgroups on their team NOTECHS scores while the vertical line has been added to indicate implementation of the **TEAM**ANATOMY program.

The X-axis represents time (measured in weeks). These charts cover pretest and posttest data points over 12 months of observational data collection. This results in 40-50 time points shown for each speciality across both phases. The number of time points aids in identifying special cause variation and in establishing a stable estimate of the underlying trend present in the pre (20-25) and post (20-25) data points [11]. The X-bar (mean) chart illustrates whether the process is stable over time [16]. Separate control charts were created to compare different specialities' performance on NOTECHS. The Y (vertical) axis is the team NOTECHS score. We calculated means (SD) of the team scores for the before-and-after time periods to assist in the interpretation of the control charts. To detect control rule violations, we used 1, 2, and 3 control limits (1SD-3SD); however, we present only the upper and lower control limits to 3SD of the mean (as is the convention). See **Appendix 1** for explanation of the control rules used in this study. Control charts were initially created using SPSS.

Results

The **TEAM**ANATOMY program training DVDs were distributed to 30 staff who were core members of the four surgical teams during the pre and post intervention phases. Of those staff, 20/30 (67%) including 4 consultant surgeons and 16 nurses viewed the DVD. Across the pretest and posttest phases, 179 surgical procedures with teams working in cardiac, vascular, upper GI, and HPB surgeries were observed. In total, 520 individual cases of observational data were collected (pretest n =292; posttest n=228). Case characteristics relative to cases observed, clinical role, length of surgery, and patient ASA are presented in **Table 1** for each surgical specialty. Across both phases, cardiac surgeries were the longest while vascular cases were the shortest. The proportion of patients with increased acuity (ASA 4) undergoing cardiac surgery was higher compared with the other specialties.

NOTECHS scores comparisons based on professional role and phases

NOTECHS scores based on professional role are shown in **Table 2**. Pretest NOTECHS revealed significant differences between groups (F=5.80, df=2/289, p=.003), using Bonferroni corrections, NOTECHs scores were similar among surgeons and anaesthetists (p=.345), however nurses differed significantly from anaesthetists only (p=.002). Furthermore, nurses also scored significantly lower in the posttest phase when compared with surgeons and anaesthetists (F=6.50, df= 2/237, p= .002). Improvements were found overall for all roles from pretest to posttest phases (Anaesthetist t=-7.7, df (171), p<.001, 95% CI -3.4 to -2.0; Surgeon t=-9.7, df (176.8), p<.001, 95% CI -4.1 to -2.7; Nurse t=-6.3, df (176), p<.001, 95% CI -3.6 to -1.9). In relation to checklist use, improvements were observed across its three phases. These differences were significant (sign-in: $\chi^2 = 6.165$, $df=1$, $p < .01$; time-out: $\chi^2 = 313.793$, $df=1$, $p < .0001$; sign-out: $\chi^2 = 67.583$, $df=1$, $p < .0001$).

Time-series analyses

Figures 1 to 4 illustrate the X-bar charts for each of the four surgical teams. The vertical dotted line in the centre of each chart indicates implementation of the training program while the horizontal line running through the middle of the chart represents the overall mean. The 99.7% control limits shown as horizontal dotted lines represent 3 SD either side of the mean. Across all teams, there appears to be improvements in the mean NOTECHS scores across the pretest-posttest phases. Visual inspection of the control chart of the cardiac team indicates that during the pretest phase, the process was mostly 'in control' with common cause variation; while in the posttest phase, there was special cause variation showing process improvement (weeks 13-20). Applying Benneyan et al's [16] control rules in the interpretation of the HPB teams' control chart, there was special cause variation across both the pretest (weeks 3-6 and 9-13) suggesting a decline in team process; while in the posttest phase (weeks 7-11 and 18) there was special cause variation with process improvement. In this team, there was also special cause variation evidenced by a posttest decline in process during weeks 19-20. Interpretation of the control chart of the upper GI team suggests pretest special cause variation (weeks 3-6 and 15, 18) indicating a decline in process, and posttest special cause variation (weeks 6-8, 10-14 and 18-20), equating to process improvement. Applying Benneyan et al's [16] control rules to the control chart of the vascular team shows pretest special cause variation (weeks 4, 18-22) indicating a decline in process, and posttest special cause variation (weeks 9-16) denoting improvements in team processes. As longitudinal analyses, these control charts illustrate process variation before-and-after implementation, which include not only the mean, but the point estimates (upper and lower control limits), and suggest that chance could be ruled out as a cause for the variation.

Discussion

Our results suggest an association between implementation of the **TEAM**ANATOMY program and post-intervention improvements in NOTECHS scores in the surgical teams observed. The observed post training improvements in NOTSS indicated special cause variation which was consistent across all teams. These results support other studies on the use of team training as an isolated intervention [18-20]. There were notable posttest improvements at across all professional groups with the most significant improvements observed in surgeons. That only 67% of study participants completed the team training program implies a lack of buy-in. At the organisational level, there will always be inherent tensions between perceptions of the potential future benefits of any safety intervention and the current costs of lost activity. Wide-scale implementation of any interprofessional training program requires a shift in the thinking of hospital administrators based on different motivators that are not solely founded on financial reward. While 2/3 of core team members participated in the training program, previous research suggests that not *all* team members require training for a 'team effect' to occur: Team culture and performance may be improved through the influence of a few [21-23]. Despite some fluidity in team membership across the teams we observed, there was on average, 2-3 core members core members per case; thus team stability may have indirectly influenced NOTSS performance

[24]. Additionally we noted significant improvements in observed adherence to the Surgical Safety Checklist procedures after implementation of **TEAM**ANATOMY. Given that the use of timeout use was targeted within the program content, this result is encouraging. Other research in this area identified positive associations between team training and NOTSS, but little evidence of impact on technical performance or patient outcome [20, 25]. We did not include clinical endpoints (e.g., post-operative complications, infection, mortality etc) as this would require much larger sample sizes to detect any significant differences in these outcomes.

Prior to developing the program, we consulted with senior medical and nursing staff to negotiate the time commitment and the most appropriate mode of delivery, given organisational constraints. Developing a one-hour DVD incorporating filmed simulations of everyday practice scenarios was innovative and responded to the contextual constraints. Staff were able to view the DVD at a time, convenient to them. Most program participants viewed the DVD in work time, during a period of reduced surgical activity. The **TEAM**ANATOMY program is evidence-informed and scenario-based, and was developed with input from senior clinicians from the study site. Program content was predicated on building individual and shared situation awareness based on cognitive, communication, teamwork and environmental factors. Based on the brevity of the program, its implementation was feasible and inexpensive.

This study makes a methodological contribution to this field: To our knowledge this is the largest prospective study of its kind in this area and the first to use the revised NOTECHS in a naturalistic setting, across a range of surgical procedures with front-line clinicians. Previously, the revised NOTECHS was tested and used only in clinical simulation.

As such it was developed primarily for training purposes where team members would have exposure to the gamut of crisis scenarios, and would consequently have opportunity to demonstrate clinical behaviours across all NOTSS factors. However in the clinical setting, surgical crises events are fortunately much less frequent. Scoring on the five items within factor five (decision-making) was problematic because very few crises were observed in real time; thus there was no opportunity for teams to demonstrate these NOTSS behaviours. Indeed, decision-making behaviours during a surgical crisis were observed in only 40-50% of cases; consequently teams posted lower scores on this factor. Therefore it was necessary to standardize individuals' scores so they were not 'penalized' through giving them lower scores, which may not truly reflect their NOTSS performance.

Although there were some challenges in using the NOTECHS tool, it has been formerly validated and its reliability was acceptable with scores remaining stable after repeated administrations [14]. The tool has certain advantages: it only requires one observer and can be used to assess the whole team and performance of subteams separately. Further as it captures NOTSS independently of other events and can be used for different types of surgeries. Our results suggest that the NOTECHS was able to detect improvements in NOTSS after specific team training across the four teams observed. While we provide control charts for each speciality we did not compare against speciality. Rather, surgical specialty was used as a grouping variable, and though differences can be seen across teams, there is no theoretical or empirical support to suggest there would be speciality differences in NOTSS performance. However, the NOTECHS provides insight and opportunity to review team and individual performance across NOTSS factors. We believe that the NOTECHS has application to a wide cross section of procedures, but recommend that further empirical work be undertaken relative to validity testing NOTECHS' behavioural dimensions to ascertain whether (or not) they actually measure the domains that reflect NOTSS. Improvements to the tool may increase its value as a research instrument.

The improvements in NOTSS performance are heartening, but there are caveats in interpreting these results. The lack of a control group to compare results against limits the extent to which they can be generalised. We considered having a control but wanted to avoid the risk of group contamination. The longitudinal design gives rise to the possibility of maturation and selection biases. For instance, 'maturation' could have occurred because surgical teams' at this facility may have developed a heightened awareness of patient safety issues, which possibly also contributed to improvements in team performance [6]. Nonetheless, the deliberate selection of teams who had worked together for at least three months was important indicated team stability and allowed per-test-post-test comparisons of the same core members. While the observer was blinded to the content of the team training program, they could not be blinded to the outcome. However across both phases, data were collected by the same single observer, enhancing consistency in the process and assessment of NOTSS. Finally as this study was based on observation, NOTSS assessment requires judgement, introducing the potential for observer bias. During instrument piloting, inter-rater checks were performed with a second trained observer, yielding an acceptable level of agreement; thus minimising the possibility of observer bias.

Conclusions

Our results suggest that implementation of a brief team training program is feasible within the constraints of a large organisation. The observed improvements in professional roles and teams' NOTSS are based on implementation of a brief team training program are encouraging. Whether the effect is sustainable over the longer term, or can be related to improved patient outcome or cost effectiveness requires further research.

Words: 3,290

References

1. Gawande A, Zinner M, Studdert D, et al. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 2003, 133:614-621.

2. Nurok M, Evans L, Lipsitz S, et al. The relationship of the emotional climate of work and threat to patient outcome in a high-volume thoracic surgery operating room team. *BMJ Qual Saf* 2011, 20:237–242.

3. Nwosu A. The horror of wrong-site surgery continues: report of two cases in a regional trauma centre in Nigeria. *Patient Saf Surg* 2015, 9:6.

4. McCulloch P, Mishra A, Handa A, et al. The effects of aviation-style non-technical skills training on technical performance and outcome in in the operating theatre. *Qual Saf Health Care* 2008, 18:109-115.

5. Greenberg C, Regenbogen S, Studdert D, et al. Patterns of communication breakdowns resulting in injury to surgical patients. *J Am Coll Surg* 2007, 204:533-540.

6. Gillespie B, Gwinner K, Chaboyer W,et al. Team communications in surgery—creating a culture of safety. *J Interprofessional Care* 2013, 27(5):287-293.

7. Armour Forse R, Bramble J, McQuillan R. Team training can improve operating room performance. *Surgery* 2011, 150:771–778.

8. Hull L, Arora S, Aggarwal R, et al. The impact of nontechnical skills on technical performance in surgery: a systematic review. *J Am Coll Surg* 2012, 214:214-230.

9. Gillespie B, Chaboyer W, Thalib L, et al. Effect of Using a Safety Checklist in Surgery on Patient Complications: A Systematic Review and Meta-Analysis. *Anaesthesiology* 2014, 120:1380-1389.

10. Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. *Qual Saf Health Care* 2009, 18:104-108.

11. Tunis S, Stryer D, Clancy C. Practical clinical trials. *JAMA* 2003, 290(12):1624-1632.

12. Gillespie B, Chaboyer W, Fairweather N. Factors that influence the expected length of operation: Results of a prospective study. *Qual Saf Health Care* 2012, 21(1):3-12.

13. Ensley M. Toward a theory of situational awareness in dynamic systems. *Hum Factors* 1995, 37(1):37-64.

14. Wright M, Endsley M. Building shared situation awareness in healthcare settings. In: *Improving healthcare team communication: Buliding on lessons from aviation and aerospace.* edn. Edited by Nemeth C. Hampshire: Ashgate; 2008: 97-114.

15. Sevdalis N, Davis R, Koutantji M, et al.Reliability of a revised NOTECHS scale for use in surgical teams. *Am J Surg* 2008, 196:184-190.

16. Benneyan J, Lloyd R, Plsek P. Statistical process control as a tool for research and healthcare improvment. *Qual Saf Health Care* 2003, 12:458-464.

17. Polit D, Chaboyer W. Statistical Process Control in Nursing Research. Res Nurs Health 2012, 35:82–93.

18. McCulloch P, Mishra A, Handa A, et al. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual Saf Heatlh Care* 2009, 18:109-115.

19. Robertson E, Hadi M, Morgan L, et al. Oxford NOTECHS II: a modified theatre team non-technical skills scoring system. *PLoS One* 2014, 9:e90320.

20. Morgan L, Hadi M, Pickering S, et al. The effect of teamwork training on team performance and clinical outcome in elective orthopaedic surgery: a controlled interrupted time series study. *BMJ Open* 2015, 5:e006216.

21. Sykes M, Gillespie B, Chaboyer W,et al. Surgical Team Mapping: Implications for Staff Allocation and Coordination. *AORN Journal* 2015, 101:238-248.

22. Salas E, DiazGranados D, J S, Weaver S, et al. Does Team Training Work? Principles for Health Care. *Acad Emerg Med* 2008, 15:1002-1009.

23. Anderson C, Talsma A. Characterizing the Structure of Operating Room Staffing Using Social Network Analysis. *Nurs Res* 2011, 60(6):378-385.

24. Gillespie B, Chaboyer W, Fairweather N. Interruptions and miscommunications in surgery: An observational study. *AORN Journal* 2012, 95(5):576-590.

25. Buljac-Samardzic M, Dekker-van Doorn C, van Wijngaarden J, et al. Interventions to improve team effectiveness: a systematic review. *Health Policy* 2010, 34:183-195.

Appendix 1

The following control rules [16] were set to detect special cause variation:

1. 1 point outside the upper or lower control limits (i.e., 3 SD).
2. 2 out of 3 successive points exceeding 2 SD from the mean on the same side as the centre line.
3. 4 out of 5 consecutive points more than 1 SD from the mean on the same side of the centre line.
4. 8 successive points on the same side as the centre line.
5. 8 consecutive points above or below the centre line.
6. 6 consecutive points increasing or decreasing (trending up or down).
7. 14 points in a row, alternating.

Table Legend

Table 1: Descriptive pretest- posttest results for case-related variables for each surgical team (n=179)

Table 2: Descriptives of NOTSS performance by professional role

Figure Legend

Figure 1: X-bar control chart for Cardiac team NOTSS

Figure 2: X-bar control chart for HPB team NOTSS

Figure 3: X-bar control chart for Upper GI team NOTSS

Figure 4: X-bar control chart for Vascular team NOTSS