Journal of Clinical Anesthesia (2016) **32**, 236–241

ELSEVIER

Journal of Clinical Anesthesia

Original contribution

Residents in tutored practice exchange groups have better medical reasoning as measured by script concordance test: a controlled, nonrandomized study

V. Compère M.D., Ph.D.[a,b], J. Abily M.D.[a,*], J. Moriceau M.D.[a], A. Gouin M.D.[a], B. Veber M.D., Ph.D.[a], H. Dupont M.D., Ph.D.[c], E. Lorne M.D.[c], J.L. Fellahi M.D., Ph.D.[d], J.L. Hanouz M.D., Ph.D.[d], J.L. Gerard M.D., Ph.D.[d], L. Sibert M.D., Ph.D.[e], B. Dureuil M.D., Ph.D.[a]

[a]*Department of Anesthesiology and Intensive Care, Rouen University Hospital, France*
[b]*Laboratory of Neuronal and Neuroendocrine Communication and Differentiation, DC2N, EA4310, U982 Inserm, Federative Institute of Multidisciplinary Research on Neuropeptides 23 (IFRMP 23), Place Emile Blondel, University of Rouen, Mont-Saint-Aignan, France*
[c]*Department of Anesthesiology and Intensive Care, Amiens University Hospital, France*
[d]*Department of Anesthesiology and Intensive Care, EA4650, Caen University Hospital, France*
[e]*Department of Medical Pedagogy, University of Rouen, Rouen, France*

Received 20 January 2015; revised 16 May 2015; accepted 5 March 2016

Keywords:
Practice exchange group;
Medical reasoning;
Script concordance test

Abstract

Study Objective: Clinical reasoning by anesthesiology residents in emergency situations where optimal management is uncertain could be improved by setting up a tutored practice exchange group. This study attempted to evaluate the impact of a practice exchange group (PEG), tutored by a senior anesthesiologist, on anesthesiology residents in emergency situations. Changes in clinical reasoning were measured by script concordance tests (SCT).

Design: We conducted a controlled, non-randomized study.

Setting and Participants: Participants are residents in anesthesiology in Rouen, Caen and Amiens University Hospitals.

Interventions: Two resident groups were made up without randomization. The first group was the control group and consisted of residents from Amiens University Hospital and Caen University Hospital. The second study group (PEG group) consisted of residents from Rouen University Hospital, who followed weekly PEG sessions. Two groups had the same learning objectives except the PEG.

Measurements: In both the control group and the study group, each resident's clinical reasoning was assessed in the same formal manner by SCT. The primary outcome measurement of this study was to compare SCT results in the study group with PEG training (PEG group) with those without (control group).

*Corresponding author: Julien Abily, Department of Anesthesia and Intensive Care, Rouen University Hospital, 1 rue de Germont, 76031 Rouen, France. Tel.: +33 2 32 88 82 83; fax: +33 2 32 88 83 26.
E-mail address: julien.abily@chu-rouen.fr (J. Abily).

http://dx.doi.org/10.1016/j.jclinane.2016.03.012

Main results: Performance in the SCT, expressed as degree of concordance with the expert panel (95% CI), was better in the PEG group (64% [62.1%-66%]) than in control group (60% [57.5%-62.8%])) (P= .004).
Conclusion: Our study strongly suggests that an expert-directed, peer-conducted educational training program may improve the clinical reasoning of anesthesiology residents as measured by SCT.

1. Introduction

The on-call period is an important component of anesthesiology residency training, because residents perform technical skills but, more importantly, they apply appropriate clinical reasoning to emergency cases, where clinical information may be incomplete and uncertainty may exist. The patient's condition may be critical and subsequent lack of time may not enable all required investigations. In addition, unstable vital parameters might make it more difficult to apply professional guidelines. Furthermore, no established recommendations are available regarding specific anesthesiology residency training in this context of emergency anesthesia care.

Peer learning or peer teaching is a method in which the student plays a direct role, not only as a teacher but also as a learner, by taking an active part in debates between peers [1]. This method has shown its efficacy in the teaching of clinical reasoning [2]. It is used particularly in practice exchange groups (PEG), also called peer or quality groups [3]. A PEG consists of a small group with same speciality physicians (peers) who meet regularly to review patient cases extracted randomly from their daily practice [4]. For each case, exchange between participants leads to a confrontation of practices. At the end of discussion, practitioners have to agree on a specific direction for action. If no satisfactory solution can be agreed upon, literature research is then initiated. In the context of continuing medical education, the purpose of these groups is to improve the quality of care, by comparison between peers and medical frames of reference [3]. This methodology has also been successfully used for initial medical education [2,5] but, to our knowledge, it has not yet been used for in anesthesiology residents who are confronted with emergency situations.

Therefore, this study attempted to evaluate the impact of anesthesiology residency training based on a PEG tutored by a senior anesthesiologist. Changes in clinical reasoning applied to anesthesia during emergency situations in a context of adult general surgery or obstetrics, were assessed by using the script concordance test (SCT).

2. Material and methods

2.1. Student selection

This prospective, controlled, non-randomized study was conducted among residents in anesthesiology from 3 French university hospitals (Rouen, Caen and Amiens). The Ethics and Evaluation Committee for Non-Interventional Research of Rouen University Hospital approved the study in April 2013 (N°E2013-9) before beginning test. All participants received information before any study procedures were undertaken. After, residents in the anesthesiology training program were invited to willingly participate as subjects in the study.

Two resident groups were made up without randomization. The first group was the control group and consisted of residents from Amiens and Caen University Hospitals. The second study group (PEG group) consisted of residents from Rouen University Hospital who have attended a weekly PEG session since 2010. Three groups had the same learning objectives except the PEG. The control group also met weekly but not with the same methodological approach and not on the same case reports. In both the control group and the study group, each resident's clinical reasoning was assessed in the same formal manner by SCT.

2.2. Study procedures

2.2.1. Learning objectives

In the 3 university centers leaning objectives for anesthesiology residents were French National Guidelines for anesthesia teaching, given to all participating residents at the beginning of their 5-year anesthesiology program.

The teaching schedule was similar for the 3 centers. This training program includes several types of learning. Students attend 1 or 2 days of specific topic lectures throughout the academic year. In addition, a weekly journal club is organized. Each centre also provides its anesthesiology residents with at least 2 training sessions on high-fidelity simulation. Assessment was common for the three universities.

2.2.2. Establishment of a tutored PEG

The PEG group of anesthesiology residents met weekly for 90 minutes at Rouen University Hospital under the supervision of a senior academic teacher. During any given session, 1 or 2 residents individually presented an anesthetic situation they had had to deal with during a recent emergency situation (obstetric case or adult emergency surgery cases). Cases were linked to difficulties associated with intensive care or daily practice. Following each case presentation, residents compared their practices in an open discussion. At the end of the exchange, the senior anesthesiologist helped when necessary to reach a final joint consensus with the approval of the entire group. The senior teacher routinely insisted on the importance of clinical reasoning based on a single diagnostic decisional

tree (in order of probability), additional examinations to be performed and medical care according to degree of severity. In addition, and when necessary, the physiology and physiopathology aspects of the case were reviewed by the group. Each student in this group had at least 8 months of PEG experience at the time of the study.

2.2.3. Assessment of clinical reasoning

The SCT confronted the residents with authentic emergency clinical situations which were described in vignettes, each of them corresponding to one of the previously set objectives. The clinical situations were problematic even for experienced clinicians, either because there were not enough data or because situations were ambiguous. There were several options for diagnosis, investigation or treatment. The items (questions) were based on a panel of questions that an experienced clinician considered relevant to this type of clinical setting. The item was consistent with the presentation of relevant options and new data (not described in the vignette). The task for the student consisted in determining the effect this new data had on the status of the option. An example of items from the therapeutic section of the test is illustrated in Table 1. All study participants had the same experience with SCT.

One independent reviewer (L Sibert) had previously expertized the study step by step. We followed the guidelines described by Lubarsky et al [6]. The SCT included initially 90 vignettes, which described clinical situations that residents might be confronted with. Each vignette included 3 items. Four authors assumed responsibility for developing items, and ensuring clarity and relevance for SCT. The first item either included a diagnostic hypothesis, plan for investigation or a treatment recommendation. Then, new information (a sign, a symptom or a result of investigation) was presented. The resident's task was to assess, using a 5-point Likert scale, the influence of this new element on the diagnostic hypothesis, the plan for investigation or treatment. The different points of the scale corresponded to positive values (the option was enhanced by the new data), neutral values (the data did not change the status of the option) or negative values (this option was ruled out by the data).

The scoring system was based on the principle that any answer given by any one expert had an intrinsic value, even if that answer did not coincide with that of the other experts [7]. According to Lubarsky et al, at least 15 anesthesiologist practitioners formed the expert panel [6]. The principles of SCT have previously been summarized by Gibot et al [8]: for each item, the answer entitled the resident to a credit corresponding to the number of experts who had chosen it. All items had the same maximum credit, and raw scores were transformed proportionally to obtain a one-point credit for the answer that was chosen by most experts. Other choices received a partial credit. Thus, to calculate the scores, all results were divided by the number of individuals who had given answers chosen by the largest number of respondents. The total score for the test was the sum of all credits earned for each item. Total score was then transformed into a percentage score. Automatic correction software was used for scoring (available at http://www.cme.umontreal.ca/tcs; September 2009). The SCT was optimized by post hoc analysis as described by Lubarsky et al [6].

The SCT was administered one time at the end of university year in June.

2.3. Outcome variables

The primary outcome was the resident's performance as measured by SCT, with and without weekly PEG. The performance in SCT was expressed as degree of concordance with the panel.

Secondary outcomes included:

- Cronbach coefficient α after optimizing test
- SCT score between training level resident and expert in each group

The data included student demographics information.

2.4. Statistical analysis

In a pilot study without PEG (conducted by the authors), including 17 residents, we previously reported a 60% (57%-63%) degree of concordance with the panel (95% CI). To calculate the sample size, we assumed that a difference of 3% between 2 groups would be pedagogically significant as recently reported by Nouh et al [9]. Based on these findings and assuming that the CI was the same between the populations and using a power of 0.90 with a level statistical significance of 0.05, it was estimated that 38 students should be included in each group. Considering the likelihood of dropouts, we enrolled 42 residents in each group.

Shapiro-Wilk was used to test for normal distribution of variables (all $P > .05$). Comparison between normally distributed variables was conducted using a parametric 1-tailed t test.

Table 1 Example of SCT applied to the case of a 20 year-old female patient admitted for a laparoscopic appendectomy for whom anesthesia must be induced

If you were considering doing:	And then you find:	The effect on the relevance of this treatment becomes:
Preoperative nasogastric tube	Fasting for 6 hours	−2 -1 0 +1 +2
Rapid sequence induction	No preoperative nausea and vomiting	−2 -1 0 +1 +2
Intravenous midazolam premedication	Preoperative anxiety	−2 -1 0 +1 +2

−2: totally or almost totally contra-indicated; −1: not useful or even detrimental; 0 neutral; +1 useful; and +2 necessary or absolutely necessary.

Variables that were not normally distributed (sub group study between training levels) were analyzed using Kruskall-Wallis test for multiple comparisons. Results which were normally distributed were expressed as means, with 95% confidence intervals. Quantitative variables which were not normally distributed were reported as median (interquartile range). Statistical analysis was performed using Prism 6.01 for windows.

3. Results

A total of 84 residents were included in the study. Flow chart is presented in Fig. 1. The PEG group was composed of 42 residents from Rouen University Hospital. The control group included 21 residents from Caen University Hospital and 21 from Amiens University Hospital. We included the residents of 3 universities with matching on the year of resident. Distribution by year of residency training was similar between the PEG study group and the control group (Table 2).

SCT including 28 scenarios for a total of 92 items was administered to a panel of 17 experts. According to recommendations by Lubarsky et al [6], we optimized SCT by performing post-hoc analysis. Items with high variability, low variability or binomial responses were excluded. We obtained a final version with 25 scenarios and 61 items. After this optimization, Cronbach coefficient α was .82.

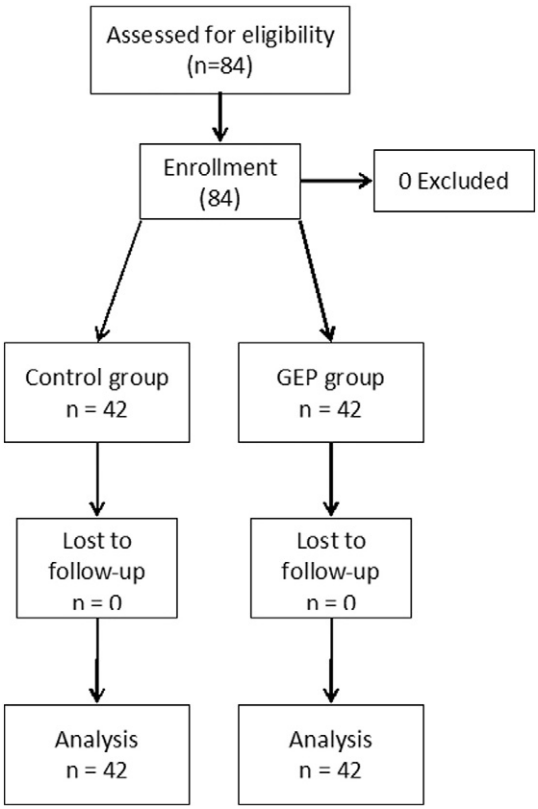

Fig. 1 Flow chart.

Table 2 Number of residents included in the study by year of training (Y)

	Control group (n = 42)	PEG group (n = 42)
First year of training (Y1) (n)	**10**	**10**
Y2 (n)	7	9
Y3 (n)	7	8
Y4 (n)	10	9
Y5 (n)	8	6

Performance in the SCT, expressed as degree of concordance with the expert panel (95% CI), was better in the PEG group (64% [62.1%-66%]) than in control group (60% [57.5%-62.8%]) (P= .004).

Fig. 2 shows tests scores by resident training year (divided into 3 different year groups: the first 2 years (RY 1 + 2; resident year 1 + 2), the third and fourth years (RY 3 + 4) and fifth year (RY5)) and expert group scores. Residents' results were significant lower than the expert group's except for the PEG RY 5 group. Difference in results between PEG RY 5 group and expert group was −8.8 (−19.45 to 1.851) expressed as mean difference (95% CI).

4. Discussion

This study showed that residents who attended PEG sessions performed better than those who did not, as measured by SCT. Furthermore, final year residents enrolled in the PEG program had the same results as the experts. This finding suggests that residents involved in the PEG program provide faster clinical

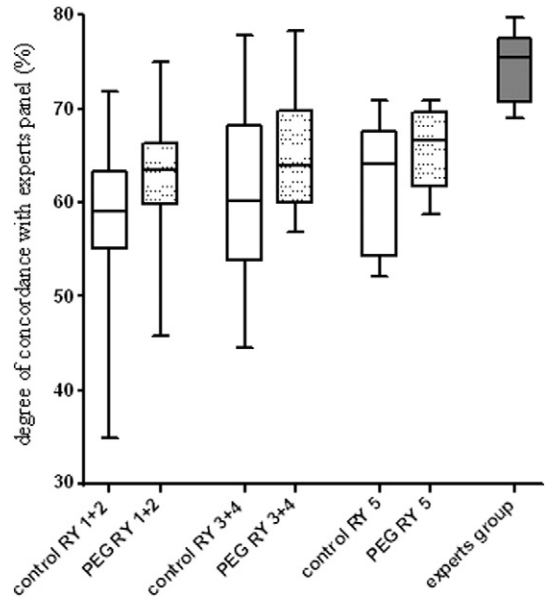

Fig. 2 Residents' scores in SCT.

reasoning, which is similar to that of experts. Nouh et al reported an increasing score measured by SCT during the resident program, which was able to differentiate junior and senior residents [9]. The same finding evidence by Park et al showing the pedagogical relevance of a difference to SCT [10]. A limitation of our study was absence of randomization. Moreover, the two groups weren't the same clinical scenarios during the weekly meeting. Many confounding variables may have an impact on student's learning. Residents learn continuously and become better during their training. Moreover, over the course of the many encounters with the residents, instructors might have influenced residents thus giving them their answers to the clinical scenario in the final examination. However, there are always several variables in educational measurement and we can postulate that the impact of these confounding parameters may be found to be equally distributed in the two groups. In fact, the PEG program has now been added to our anesthesiology residency training at Rouen University Hospital and may reflect a gain in clinical practice over this period. In addition, there was no pre-test, and this should be considered in the design of future studies.

The emergency situations are difficult to manage in real time, even for an experienced practitioner, because all appropriate means for anesthetic management must be considered [11]. This difficulty raises the critical issue of the initial education of residents who are confronted with such uncertain situations. Our study suggests that the establishment of PEG could improve residents' clinical reasoning performances as measured by SCT. In fact, this particular form of interactive teaching allows residents to play a direct role, not only in their own education but also, in the training of other learners actively participating in discussions between peers [1]. According to the literature, providing resident with teaching responsibilities is associated with an improvement in clinical skills and clinical reasoning [12–14]. PEG programs are part of the anesthesia-intensive care curriculum and teach ethical reasoning applied to residents' training [15]. Perception of the benefit of the teaching sessions assessed using a structured questionnaire was high (always $\geq 8/10$). This indicated that the PEG learning sessions were considered as an integral part of the overall educational process, enabling management of ethical issues relevant to the practice of anesthesia and critical care. Such discussions also encouraged dialogue between the people involved and residents.

In a specialty where the constant concern is ensuring patient safety, other tools for training and evaluating residents in management of risk situations have previously been assessed. For example, simulation is a discriminating and valid tool for evaluating residents' performance in critical situations [16,17]. However, this main disadvantage of simulation is cost, not only financially but also in human resources. Another method of clinical reasoning that could be used is training based on problem solving. This method has now been developed for residents [18,19] and Yazigi et al have reported interesting results in the field of anesthesia [20]. Not only learners, but also tutors reported positive educational benefit from this training method. However, the real impact on knowledge acquisition and thinking among residents, as well as development of their clinical competences, were not studied in depth.

In this study the construction of SCT followed academic methodology. According to Lubarsky et al, a small number of authors must participate in the construction and review of items, inspired by situations in their daily practice [6]. We included four anesthesiologists, involved in anesthesiology teaching, to develop SCT. In our work, 92 items divided into 28 clinical cases were then given to a panel of 17 experts. Gagnon et al have shown that at least 15 panel members are required to obtain adequate reliability of scores, and only marginal benefit is gained by having more than 20 panel members [21]. The choice of experts can be problematic and must be consistent, taking into account those competences which are to be assessed. In the present study, the expert panel of anesthesiologists was involved in emergency anesthesia and critical care for adult general surgery and/or obstetrics. It is obvious that since SCT is a written examination, it does not, allow assessment all areas of clinical competence such as data collecting during the interview or physical examination. Following the first step, post-hoc analysis for optimization was performed. Items with high variability, low variability or binomial responses were excluded. Approximately 25% of items are generally excluded [22]. In our case, 92 items were initially constructed and only 61 were finally selected. We obtained a final Cronbach score of 0.82. In evaluation practice, a test is reliable when this ratio has a value greater than 0.80 [23]. The main disadvantage of SCT is cost in human resource for the construction of tool. After their validations, the SCT could be evolved easily and cheaply.

In conclusion, this study strongly suggests that an expert-directed, peer-conducted educational training program may improve the clinical reasoning of anesthesiology residents as measured by SCT.

Acknowledgement

The authors are grateful to Nikki Sabourin-Gibbs, Rouen University Hospital, for her help in editing the manuscript.

References

[1] Walker-Bartnick LA, Berger JH, Kappelman MM. A model for peer tutoring in the medical school setting. J Med Educ 1984;59(4):309-15.
[2] Scherpbier A, Busari J. Why residents should teach: A literature review. J Postgrad Med 2004;50(3):205-10.
[3] Santé HAS. Le groupe d'analyse des pratiques entre pairs—"peer preview". Haute Autorité de Santé; 2006.
[4] Cohendet C, Jacot P, Arnould M, Bertier A, Gorel J-L, Mas J. Groupes de pairs: l'évaluation et la formation professionnelle au diapason de l'EBM. EBM J Éd Fr 2003;34:1-2.
[5] Secomb J. A systematic review of peer teaching and learning in clinical education. J Clin Nurs 2008;17(6):703-16.

[6] Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: From theory to practice: AMEE Guide No. 75. Med Teach 2013;35(3):184-93.
[7] Charlin B, Boshuizen H, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. Med Educ 2007;41(12):1178-84.
[8] Gibot S, Bollaert P-E. Le test de concordance de script comme outil d'évaluation formative en réanimation médicale. Pédagog Méd 2008; 9(1):7-18.
[9] Nouh T, Boutros M, Gagnon R, Reid S, Leslie K, Pace D, et al. The script concordance test as a measure of clinical reasoning: a national validation study. Am J Surg 2012;203(4):530-4.
[10] Park AJ, Barber MD, Bent AE, Dooley YT, Dancz C, Sutkin G, et al. Assessment of intraoperative judgment during gynecologic surgery using the Script Concordance Test. Am J Obstet Gynecol 2010;203(3):240.e1-6.
[11] Perrenoud. Enseigner: agir dans l'urgence, décider dans l'incertitude. Savoirs et compétences dans un métier complexe. Paris: ESF; 1996.
[12] Sheets KJ, Hankin FM, Schwenk TL. Preparing surgery house officers for their teaching role. Am J Surg 1991;161(4):443-9.
[13] Minor S, Poenaru D. The in-house education of clinical clerks in surgery and the role of housestaff. Am J Surg 2002;184(5):471-5.
[14] Pelletier M, Belliveau P. Role of surgical residents in undergraduate surgical education. Can J Surg 1999;42(6):451-6.
[15] Yazigi A, Madi-Jebara S, Yazbeck P. Les séances d'apprentissage du raisonnement éthique appliquées à la formation des résidents en anesthésie-réanimation. Pédagog Méd 2007;8(3):177-83.
[16] Lebuffe G, Plateau S, Tytgat H, Vallet B, Scherpereel P. Interest of mannequin based simulator to evaluate anaesthesia residents; 2005 260.
[17] Gaba DM, Howard SK, Flanagan B, Smith BE, Fish KJ, Botney R. Assessment of clinical performance during simulated crises using both technical and behavioral ratings. Anesthesiology 1998;89(1): 8-18.
[18] Chamberland M. Clinical reasoning learning (CRL) sessions. An example of a contextualized teaching activity adapted to clinical stages in medicine; 1998 479.
[19] Philp E, Parker P, Rubin N, Tietze P, Philp J. The difficult patient: creation of a curriculum by third-year family practice residents. Fam Med 1996;28(8):553.
[20] Yazigi A, Nemr E, Jaoudé SA, Antakly M-C. L'apprentissage par la méthode des cas appliqués à la formation des résidents en anesthésie-réanimation. Pédagog Méd 2002;3(4):232-6.
[21] Gagnon R, Charlin B, Coletti M, Sauve E, van der Vleuten C. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? Med Educ 2005; 39(3):284-91.
[22] Dory V, Gagnon R, Vanpee D, Charlin B. How to construct and implement script concordance tests: insights from a systematic review. Med Educ 2012;46(6):552-63.
[23] Charlin B, Gagnon R, Sibert L, Van der Vleuten C. Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. Pédagog Méd 2002;3(3):135-44.