

Reliability Of A Peer-Approved Checklist for Evaluation of Medical Students' Anesthesia Reports

Polpun Boonmak, MD, Suhattaya Boonmak, MD, Duenpen Horatanaruang, MD and Piyaporn Bunsangjaroen, MD

*DEPARTMENT OF ANESTHESIOLOGY, FACULTY OF MEDICINE, KHON KAEN UNIVERSITY, KHON KAEN, 40002, THAILAND

Abstract

Objective: To study the reliability of a peer-approved checklist for scoring anesthesia case reports prepared by 5th-year medical students.

Methods: A report checklist was created by a tutor in the Department of Anesthesia, Khon Kaen University, comprising: 4 items on pre-anesthetic evaluation; 4 on pre-anesthetic preparation; 8 on intra-operative management; and, 5 on postoperative care. The checklist was pre-approved by all 14 anesthesiology tutors in the Department. Each student prepared case report was evaluated by 4 tutors (in random order) using the checklist. The reports and checklist were collected for inter-rater and intra-class correlation analyses.

Results: Thirty anesthesia case reports were included. The kappa (**K**) coefficient for each item was between 0.10 and 0.96: 3 items (14.28 % of the total items) had almost perfect reliability (**K** = 0.81-1.00); 2 (9.53 %) substantial (**K** = 0.61-0.80); 3 (14.28 %) moderate (**K** = 0.41-0.60); 4 (23.81 %) fair (**K** = 0.21-0.40); and, 12 (38.10 %) slight (**K** = 0.0-0.20). Intra-class correlation coefficients within each part were between 0.51 and 0.70. The whole anesthesia report had an intra-class correlation of 0.75.

Conclusion: Reliability of the anesthesia scoring checklist for the case reports of 5th-year medical students was low, despite improvements in the scoring checklist. Scores from case report writing may not reflect either the knowledge or skill levels of medical students vis-à-vis their anesthesia patients. The checklist for case report scoring should be further tested for reliability and improved before using.

Keywords: Anesthesia Reports; Reliability; Medical Students; Scoring

Background

Medical case report writing is a necessary part of medical education because it demonstrates knowledge and synthesis to convey experience to others¹. It is, therefore, used to evaluate thoroughness of understanding and skills competency among medical students.

Medical case reports are evaluated by the attending tutor who knows the patient. Sometimes the report score depends on each tutor's judgment. The Department of Anesthesiology, Khon Kaen University, used to score its medical students' case reports by giving points for: pre-anesthetic evaluation and preparation, anesthetic management, postoperative care and problem management, but without providing any pre-defined criteria or checklist for each part. Expectedly, anesthesia case reports score for 5th-year medical students varied greatly depending and grading was necessarily subjective.

With an appropriate evaluation of case reports, anesthesia case report writing could become a valid tool for student assessment. A number of studies indicate that checklists usually have low validity and poor reliability.²⁻³ Our Department therefore developed a peer-approved checklist for scoring the anesthesia case reports of 5th-year medical students.

Objectives

To study the reliability of a peer-approved checklist for scoring the anesthesia case reports of 5th-year medical students.

Material and method

Subjects

Thirty anesthesia case reports prepared by 5th-year medical students.

Method

The Ethics Committee, Khon Kaen University, approved the study protocol. A report checklist was created by tutors in the Anesthesia Department, Faculty of Medicine, Khon Kaen University, comprised 4 parts (total score = 100): 4 items in the first part (pre-anesthetic evaluation), 4 in the second part (pre-anesthetic preparation), 8 in the third part (intra-operative management) and 5 in the fourth part (post-operative care) (Table 1). The checklist was approved by all 14 anesthesiology tutors in the Department before use.

Each case report was evaluated using the checklist by 4 tutors (in random order). None of the tutors were aware of the score others gave. Then the tutors scored according to the checklist. All the reports and checklists were collected by the author, and the scores for each item from each tutor recorded.

Statistical analysis

The score for each item was analyzed for its inter-rater agreement (weighted κ coefficient). The scores for each part and the whole report were analyzed for intra-class correlations.

Table 1: Medical student's anesthesia report checklist

Part	Items	Excellent	Good	Fair	Poor	Factor
1. Pre-anesthetic evaluation	History taking					1
	Physical examination					1
	Medical problems conclusion					2
	ASA classification evaluation	correct	-	wrong	-	1
2. Pre-anesthetic preparation	Informed patient					1
	choice of anesthesia and reason					2
	Investigation					1
	NPO time	correct	-	wrong	-	1
3. Intra-operative management	Equipment preparation					1
	Drug and dose					2
	Sequence of GA/RA					2
	Monitoring during anesthesia					1
	Complication expectation					2
	Evaluation of extubation (only in general anesthesia report)					0.5
	Intravenous fluid management	correct	-	wrong	-	1
Transferring from operating room					0.5	
4. Post-operative care	Postoperative monitoring					1
	Complication					1
	Postoperative pain control					1
	Postoperative care					1
	PAR score	correct	-	wrong	-	1

Results

Table 2: Inter-rater weighted κ and Intra-class correlation coefficients for each part

Part	Items	Inter-rater weighted κ	Intra-class correlation(95% CI)
1. Pre-anesthetic evaluation	History taking	0.22	0.51(0.32-0.69)
	Physical examination	0.27	
	Medical problems conclusion	0.10	
	ASA classification evaluation	0.77	
2. Pre-anesthetic preparation	Informed patient	0.44	0.56(0.39-0.73)
	choice of anesthesia and reason	0.13	
	Investigation	0.27	
	NPO time	0.77	
3. Intra-operative management	Equipment preparation	0.18	0.70(0.83-0.95)
	Drug and dose	0.56	
	Sequence of GA/RA	0.15	
	Monitoring during anesthesia	0.48	
	Complication expectation	0.20	
	Evaluation of extubation (only in general anesthesia report)	0.96	
	Intravenous fluid management	0.91	
	Transferring from operating room	0.20	
4. Post-operative care	Postoperative monitoring	0.14	0.70(0.55-0.82)
	Complication	0.23	
	Postoperative pain control	0.34	
	Postoperative care	0.12	
	PAR score	0.96	
Report score			0.75(0.61-0.85)

The **K** coefficient for each item was between 0.10 and 0.96 (Table 2). The intra-class correlation coefficient within each part was between 0.51 and 0.70. The whole anesthesia case report checklist had an intra-class correlation of 0.75 (Table 2).

Discussion

In our study, the **K** coefficient varied from 'very high' to 'very low' (0.10-0.96) reflecting the variation in the reliability of the report checklist: three (14.28% of the total) had almost perfect reliability⁴ (**K**, 0.81-1.00); two (9.53%) substantial reliability (**K**, 0.61-0.80); three (14.28%) moderate reliability (**K**, 0.41-0.60); four (23.81%) fair reliability (**K**, 0.21-0.40); and, 12 (38.10%) slight reliability (**K**, 0-0.20). Almost all the items that gave a high reliability were 'correct/wrong' items on the checklist; while items with the poorest reliability were items with a 4-grade checklist. Our checklist for scoring anesthesia case reports had proven content-validity, based on a peer review by all of the tutors in the Department. Notwithstanding, inter-rater agreement for some items was low especially for rated items. Contrary to some studies, where the global rating scale scored by the experts provided better inter-station reliability, construct- and concurrent-validity, than did checklists^{5,6}. Some reports found more reliability on the checklist than the global rating scale for an assessment of construct validity⁷⁻⁹. Inter-observer variability was similar, whether a checklist or global assessment rating scale was used¹⁰.

The possible reasons for the poor reliability in our study may be the quality of the checklist items or the understanding and interpretation of each item by the tutors. In spite of the poor reliability of many of the items, intra-class correlation of each part and of the whole case report are moderately acceptable. So a score based on the whole report could still be reliable for student evaluations.

A valid and reliable checklist can be an accurate tool for evaluating students. Ogdan GR, et al. even used students as examiners in OSCEs¹¹, and MacRae HM, et al. have reported that checklist scores correlated strongly with ratings by physicians of history taking and physical-examination skills, providing some evidence of their validity¹². The large number of checklist items on our marking sheet might be a problem. Wilkinson TJ, et al. found an increase in number of checklist items decreased both reliability and validity, and achieving objectivity requires diligent examiners involved in the whole assessment¹³.

Conclusion

Reliability of a peer-approved checklist for scoring anesthesia case reports of 5th-year medical students was low. Scores from case report writing may therefore not accurately reflect knowledge and skills of medical students of their anesthesia subjects. Multiple factors may have affected the reliability of our checklist for scoring anesthesia case report. Improving some items with poor reliability and improving tutors' understanding of each item might make the checklist more reliable.

Acknowledgments

The authors thank Mr. Bryan Roderick Hamman for assistance with the English-language presentation of the manuscript.

References

1. Thomas R, Nieman LZ, Holbert D. The medical record and the medical interview: an evaluation of student case histories. *Fam Med* 1987 ;19(6):449-52.
2. Gorter S, Rethans JJ, Scherpbier A, van der Heijde D, Houben H, van der Vleuten C, van der Linden S. Developing case-specific checklists for standardized-patient-based assessments in internal medicine: A review of the literature. *Acad Med* 2000 ;75(11):1130-7.
3. Travis TA, Colliver JA, Robbs RS, Barnhart AJ, Barrows HS, Giannone L, et al. Validity of simple approach to scoring and standard setting for standardized-patient cases in an examination of clinical competence. *Acad Med* 1996;71(1):S84-6.
4. Hull CH. Kappa inter-rater agreement. In: Hull CH ed. *Stata reference manuals release 6*, Texas: College Station; 1999: 132-43.
5. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 1998;73(9):993-7.
6. Kramer AW, Jansen JJ, Zuithoff P, Dusman H, Tan LH, Grol RP, Van der Vleuten CP. Predictive validity of a written knowledge test of skills for an OSCE in postgraduate training for general practice. *Med Educ* 2002 ;36(9):812-9.
7. Goff BA, Lentz GM, Lee D, Houmard B, Mandel LS. Development of an objective structured assessment of technical skills for obstetric and gynecology residents. *Obstet Gynecol* 2000 ;96(1):146-50.
8. Hemmer PA, Hawkins R, Jackson JL, Pangaro LN. Assessing how well three evaluation methods detect deficiencies in medical students' professionalism in two settings of an internal medicine clerkship. *Acad Med* 2000 ;75(2):167-73.
9. McCurdy FA, Weisman LE. Teaching newborn medicine to third-year medical students. Use of a checklist. *Arch Pediatr Adolesc Med* 1995 ;149(1):49-52.
10. LaMantia J, Rennie W, Risucci DA, Cydulka R, Spillane L, Graff L, et al. Inter-observer variability among faculty in evaluations of residents' clinical skills. *Acad Emerg Med* 1999 ;6(1):38-44.
11. Ogden GR, Green M, Ker JS. The use of interprofessional peer examiners in an objective structured clinical examination: can dental students act as examiners? *Br Dent J* 2000 ;189(3):160-4.
12. MacRae HM, Vu NV, Graham B, Word-Sims M, Colliver JA, Robbs RS. Comparing checklists and databases with physicians' ratings as measures of students' history and physical-examination skills. *Acad Med* 1995 ;70(4):313-7.
13. Wilkinson TJ, Frampton CM, Thompson-Fawcett M, Egan T. Objectivity in objective structured clinical examinations: checklists are no substitute for examiner commitment. *Acad Med* 2003 ;78(2):219-23.