European Archives of Oto-Rhino-Laryngology (2019) 276:383–388
https://doi.org/10.1007/s00405-018-05268-6

OTOLOGY

Rating surgical field quality in endoscopic ear surgery: proposal and validation of the "Modena Bleeding Score"

Matteo Alicandri-Ciufelli[1,2] · Luca Pingani[3] · Davide Mariano[1] · Lukas Anschuetz[4] · Giulia Molinari[1] · Daniele Marchioni[5] · Marco Bonali[1] · Gian Maria Galeazzi[6] · Livio Presutti[1]

Received: 24 August 2018 / Accepted: 24 December 2018 / Published online: 2 January 2019

Abstract

Purpose To develop and validate a bleeding score that could be applied in endoscopic ear surgery (EEarS).

Methods A prospective validation study was performed. A new bleeding score, called "Modena Bleeding Score" (MBS), was created by the authors. It provides five grades for rating the surgical field during EEarS procedures (from grade 1—no bleeding to grade 5—bleeding that prevents every surgical procedure except those dedicated to bleeding control). A preliminary "face validity" was performed by 18 ENT specialists to assess possible misunderstandings in interpreting the scale. Then, 15 videos of endoscopic ear surgery procedures, each divided into three parts ($t0$, $t1$, and $t2$), were subsequently evaluated by 15 specialists, using MBS. The videos were randomly selected and assigned. Intra-rater reliability and inter-rater reliability were calculated. The clinical validity of the instrument was calculated using a referent standard (i.e., four ENT experts whose ratings were compared to those obtained by the former sample).

Results The face validity showed a good consensus about the clarity and comprehension of the scale; both intra and inter-rater reliability demonstrated good performance (intra-rater reliability ranged from 0.741 to 0.991 and inter-rater reliability was 0.790); clinical validity also showed positive values, ranging from 0.75 to 0.93.

Conclusions MBS has proved to be an effective method to rate surgical field during EEarS, with good-to-excellent performances. Its use would possibly help comparisons of groups in clinical trials or comparisons between studies.

Keywords Endoscopic ear surgery · Modena Bleeding Score · Surgical field rating · Endoscopic surgery · Middle ear surgery · Bleeding

Introduction

Nowadays, endoscopy is applied to a wide range of fields, both for diagnostic and interventional purposes. One of the most significant advantages of endoscopy is the direct and clear visualization of anatomical structures during the entire procedure. Anyway in endoscopic techniques, the tip of the endoscope lies directly inside the surgical field. This can lead to blood or debris adhering to the endoscope with subsequent deterioration of vision.

Bleeding, in particular, could seriously impair the quality of the surgical field. Bleeding control is even more crucial in small highly vascular cavities, such as the paranasal sinuses or the joints, where very small amounts of blood can lead to the termination of surgery. Uncontrolled intraoperative bleeding during endoscopy may result in a poor view of anatomical landmarks, prolonged operative time, and threat of surgical performance with possible higher rate of

✉ Giulia Molinari
dr.giuliamolinari@gmail.com

1 Otolaryngology-Head and Neck Surgery Department, University Hospital of Modena, Via del Pozzo 71, 41125 Modena, Italy

2 Neurosurgery Department, New Civil Hospital Sant'Agostino-Estense, Stradello Baggiovara 53, 41126 Baggiovara, MO, Italy

3 Human Resources, Department of Mental Health, Local Health Agency Reggio Emilia, Reggio Emilia, Italy

4 Department of Otorhinolaryngology, Head and Neck Surgery, Inselspital, Bern University Hospital, Bern, Switzerland

5 Otolaryngology-Head and Neck Surgery Department, University Hospital of Verona, Piazzale Aristide Stefani 1, 37126 Verona, Italy

6 Psychiatry Department, University of Modena and Reggio Emilia, Via del Pozzo 71, 41100 Modena, Italy

complications, as shown by the experience of many authors in endoscopic sinus surgery (ESinS) [1–4].

The clarity of surgical field is, hence, very important and a number of bleeding control methods, in different surgical scenarios, are available. Nevertheless, their comparison in terms of effectiveness is complex, also because, to present authors' knowledge, validated and universally accepted methods to assess the quality of the surgical field and quantify bleeding are not widespread.

Endoscopic instrumentation, techniques, and knowledge have improved during the last few years, and we believe that, in the future, endoscopic surgical techniques will increase their importance in otologic surgery. Factors considered crucial drawbacks in endoscopic ear surgery (EEarS) are the fact that is a one-handed technique, and that bleeding control in such a small space (i.e., tympanic cavity) could be very challenging.

The Fromme–Boezaart grading scale is a well-known method to assess intraoperative bleeding, which was validated for ESinS only [5, 6]. The drawback of Fromme–Boezaart scale is that it depends on the frequency of suctioning to cleanse the surgical field, so it could not be applied in surgeries where suction instruments are used to elevate or dissect tissues. In addition, the Wormald scale was created to assess bleeding, but with a marked specificity for ESinS, since it is mainly based on the time required to fill the sphenoid sinus [6].

At the present time, neither specific nor validated methods of rating quality of the surgical field in EEarS are available.

The aim of this paper would be to develop, describe, and validate a bleeding scoring system for EEarS, named "Modena Bleeding Score" (MBS). The use of a uniform and validated bleeding score could facilitate the comparison of results between groups in trials or of results from different studies.

Methods

The Modena Bleeding Score (MBS)

The MBS is a categorical scale written in the English language, developed to estimate the quality of the surgical field in relation to bleeding during the surgical procedure. It is composed by five different levels (from "Grade 1—no bleeding" to "Grade 5—bleeding that prevents every surgical procedure except those dedicated to bleeding control") (Table 1).

Face validity

To assess the face validity, the MBS was presented to a sample of surgeons from the Otorhinolaryngology Department of the University Hospital of Modena (Italy). After a 2-h training, all the surgeons were asked to answer to the following questions for each score level: (1) Is this the level that took longer to be understood? (2) Is this level expressed in a clear and comprehensible way? (3) Do you think this level can be expressed more clearly?

In addition, some personal information was requested (socio-demographic data and information about the activity performed as a surgeon).

Intra-rater and inter-rater reliability

Fifteen surgeons currently working at the Otorhinolaryngology Department of the University Hospital of Modena gave their approval to be involved in the study as assessors for measuring intra-rater and inter-rater reliability. Fifteen videos were randomly selected from the archive of the video recordings of endoscopic ear surgical procedures performed at our department. In each video, three bleeds were detected ($t0$, $t1$, and $t2$). Considering that each evaluator watched three video recordings, each of which containing three bleeding, a number of evaluators equal to 15 were sufficient for a study of agreement relative to a scale of five-item rating [7].

Each evaluator was asked to evaluate the same randomly selected video twice, at 15-day distance using the MBS (intra-rater reliability). Each evaluator was also asked to fill out the questionnaire in relation to three other videos: the evaluations thus obtained were subsequently compared to those of the other evaluators (inter-rater reliability). Intra-rater reliability was calculated using Spearman's rank correlation coefficient ranging from 1 (perfect negative correlation) to 1 (perfect positive correlation), while Intraclass

Table 1 Modena Bleeding Score

	Scoring
No bleeding	1
Bleeding easily controlled by suctioning, washing, or packing without any significant modification or slowing of surgical procedure	2
Bleeding slowing surgical procedure	3
Most of the maneuvers dedicated to bleeding control	4
Bleeding that prevents every surgical procedure except those dedicated to bleeding control	5

Correlation Coefficient [8, 9] was used for calculating inter-rater reliability (less than 0.40: poor; between 0.40 and 0.59: fair; between 0.60 and 0.74: good; between 0.75 and 1.00: excellent).

Clinical validity

The clinical validity of the MBS was calculated using a gold standard. A group of four medical specialists in otorhinolaryngology (not involved in other areas of this study) collegially viewed and evaluated all the 45 bleedings present in the 15 selected videos. After extensive discussion, the group defined a unanimous score through the MBS for each bleeding (to be referred to as referent standard). The referent standard evaluations were then compared with those obtained for inter-rater reliability. The agreement level was calculated through the Cohen "K" [10] (< 0 no agreement; 0–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; 0.81–1 perfect agreement).

Results

Face validity

The sample was composed by 18 ENT surgeons working at the University Hospital of Modena: all participants attended the training and answered the questionnaire related to the examinations of the face validity. As described in Table 2, the 61.1% ($N = 11$) of the sample was composed by males, while the most represented age group was between 26 and 35 years. Six surgeons (33.3%) had already used a questionnaire for the evaluation of bleeding. Of all participants, 55.6% ($N = 10$) read more than 30 scientific papers per year and only two surgeons (11.1%) had been working for more than 20 years at the Otorhinolaryngology Department. The average time used to complete the questionnaire was 34.2 s (DS = ± 11.6), with a minimum value of 20 and a maximum of 60 s.

Table 3 shows the values obtained from the scale in terms of clarity and understanding. The scale item that took more time for understanding was the Grade 2 ($N = 8$; 44.4%). The consensus about the clarity and comprehension of the various items was great for item 1 (100%) and extremely positive for the other items (94.4%). For items 2, 4, and 5, suggestions were also made for possible changes to be made, as described in Table 3.

The intra-rater reliability analysis gave positive and encouraging results: Spearman's rank correlation coefficient values were all greater than 0.700 (ranging from 0.741 to 0.991) and statistically significant present for all three evaluations ($t0$, $t1$, and $t2$) (Table 4). The inter-rater reliability was excellent as the coefficient is equal to 0.790 (Table 5).

Table 2 Socio-demographic characteristics of the sample and information about the experience in surgical clinical setting

	Median	*N*	%
Age			
26–35	26–35	11	61.1
36–45		4	22.2
46–55		2	11.1
56–65		1	5.6
Sex			
Male	nc	11	61.1
Female		7	38.9
Have you already used questionnaires for the assessment of bleeding?			
Yes	nc	6	33.3
No		12	66.7
Scientific articles read in 1 year concerning ENT surgery			
6–10	> 30	1	5.6
11–15		0	0
16–20		3	16.7
21–25		2	11.1
26–30		2	11.1
> 30		10	55.6

Clinical validity

The clinical validity of the instrument was extremely positive and encouraging: $K = 0.75$ for bleeding at $t0$, $K = 0.93$ at $t1$, and $K = 0.85$ for bleeding at $t2$.

Discussion

Surgical field quality and good visualization are considered crucial points for surgeons, and studies that compare interventions to improve those factors are present in the literature [1–4]. For this reasons, methods for quantifying surgical field quality have been proposed in the past.

Bleeding is most probably the main factor that can influence visualization during endoscopic procedures. Bleeding control is even more crucial in small and highly vascular cavities, where even very little bleeding can threaten the continuation of surgery. For example several studies have been carried out to evaluate the impact of anesthesia (anesthetic drugs, types of anesthesia, and ventilation modalities) on the quality of surgical field and bleeding [11–15]. Most of them were performed in endoscopic sinus surgery (ESinS), where issues related to small room and one-handed

Table 3 Analysis of face validity

	Is this the level that took longer to be understood?	Is this level expressed in a clear and comprehensible way?	Do you think this level can be expressed more clearly?	Recommendations to increase the clarity of the scale
Grade 1	0	Yes: 18 (100%)	No: 18 (100%)	None
Grade 2	8 (44.4%)	Yes: 17 (94.4%)	Yes: 3 (16.7%)	Easily controlled, not slowing surgical procedure
Grade 3	1 (5.6%)	Yes: 17 (94.4%)	Yes: 1 (5.6%)	None
Grade 4	1 (5.6%)	Yes: 17 (94.4%)	Yes: 2 (11.1%)	Add the following sentence: "Change of surgical strategy due to bleeding"
Grade 5	7 (38.9%)	Yes: 17 (94.4%)	Yes: 5 (27.8%)	Add the following sentence: "Bleeding causing interruption of the surgical procedure" Impossibility to proceed with surgical steps Bleeding makes any surgical procedure impossible

Table 4 Intra-rater reliability

	First video view		
	*t*0	*t*1	*t*2
Second video view			
*t*0	ϱ = 0.741 p = 0.002		
*t*1		ϱ = 0.991 p < 0.001	
*t*2			ϱ = 0.788 p < 0.001

maneuvering are present, as in EEarS. The methods used to quantify bleeding during ESinS can be broadly divided into two main categories, namely objective quantifications and subjective quantifications:

- An objective quantification is based on the measurement of blood loss during the surgical procedure. For example, a method for assessing blood loss consists in measuring the volume of suctioned fluids [13, 14]. Beule et al. have described a method, which compares the preoperative hemoglobin (Hb) in patient's blood with concentration of Hb in the suction unit. Although these methods employ objective parameters, they have the disadvantage of not being easily accessible for immediate use, due to their duration, expensiveness, and requirement of specific laboratory for analysis. Moreover, they do not consider the amount of blood ingested by the patient and do not take into account the irrigation solution. These procedures can give only an overall impression of bleeding, describing the total amount of blood loss at the end of surgery, and are not able to describe bleeding during procedures time by time.
- A subjective quantification is based on scales for assessing surgical field visibility. Those methods give a measure of bleeding during the surgical time, making them dynamic instruments. For example, the visual analog scale (VAS) has been used in several studies [13–15]. VAS is a simple 10-point numeric rating scale, ranging from 0 (namely no visual impediment) to 10 (worst vision). The most important limitation in using a VAS in this very topic is that the individual interpretation of the scale prevents a strong inter-rater reliability [16]. The Fromme–Boezaart scale is a six-point scale based on the frequency of suctioning required to maintain clear visibility of the surgical field [5] (e.g., Grade 1 Slight bleeding–no suctioning of blood required or Grade 4 Moderate bleeding—frequent suctioning required. Bleeding threatens surgical field directly after suction is removed). Fromme–Boezaart scale was validated on ESinS, and to present authors' knowledge, this is the only validated scoring system for surgical field in the literature [6]. Drawback of this scale is its dependence on a single action (i.e., suctioning) for rating. In EEarS suction instruments are used very frequently for dissecting/elevating tissues, and this could potentially alter the rating. Moreover, some sets of suctioning instruments for dissection in EEarS are becoming popular (e.g., Panetti set, Spiggle and Theis, Germany) and this would make the Fromme–Boezaart unsuitable for EEarS. The Wor-

Table 5 Inter-rater reliability

	Intraclass correlation	95% Confidence interval		*F* test with true value 0			
		Lower bound	Upper bound	Value	*df*1	*Df*2	*p*
Single measures	0.790	0.502	0.924	4.77	14	28	≤ 0.001

mald scale is a 11-point scale, based on number of oozing points as well as on the time required to fill the sphenoid sinus [6]. Size of the sphenoid and its ostium can impact on filling time, so the scale can be less precise. Moreover, this scale refers to a specific anatomical structure used to quantify bleeding (sphenoid sinus), and this would make inapplicable to the other surgical fields, similarly to the Fromme-Boezaart.

For the above-mentioned reasons, we decided to introduce and validate a bleeding score that could fit to EEarS, independently from a specific instrument or anatomical structure. In our study, the MBS has shown to be immediately clear, in fact, at the face validity, no evaluator needed more than 1 min to read and completely understand the scale. It was particularly important that each level of the scale has received a close unanimous favorable impression about clarity and comprehension (four out of five items were of great positive consensus). The analysis carried out on a representative sample (in which 66% of participants have not previously used questionnaires to grading surgical bleeding) has shown that the point two of the scale was the most difficult to be understood. This may be interpreted as difficulty to define the meaning of "slowing" surgical procedures. Moreover, the definition "slowing" was susceptible to subjective judgment. Although this mentioned limitation observed during the face validity, a greater uniformity of opinion in scoring by evaluation of the videos has been observed.

About performances of MBS, both intra-rater (Table 4) inter-rater (Table 5) reliability gave positive results. The intra-rater reliability ranged from 0.741 to 0.991, while the inter-rater reliability was excellent: 0.790. Equally encouraging result was obtained in clinical validity; the comparison between inter-rater reliability and the referent standard (referred to the group that collegially evaluated all videos during clinical validity phase) produced a range of values from 0.75 to 0.93, corresponding to an agreement level from substantial to perfect.

Like other scales, the MBS is able to give only a subjective evaluation of bleeding during surgery. As previously emphasized, a qualitative measure better applies in most cases to this purpose, and are, at the present time, the most popular type of grading. Although the qualitative evaluation performed by the mentioned scales does not give an accurate quantification of blood loss during surgical procedures, they are dynamic and easy instruments that enable the rater to quickly compare different surgical scenarios.

Present authors deem that the independence of the MBS from suctioning during surgery represents a significant advantage. In fact, suction is also used for performing some surgical steps, particularly during the elevation of the tympano-meatal flap. Moreover, as mentioned above, elevators with integrated suction cannulae are becoming popular, and at the present time, bleeding control could be achieved in several ways: hemostatic drugs, hemostatic devices, technical actions (such as packing or waiting), or different anesthetic strategies.

Eventually in the future, the effectiveness and comparison of different methods for achieving bleeding control in the context of EEarS could be better investigated applying the MBS. A further possible advantage is that nothing uniquely and exclusively related to a surgical specialty is present in the MBS. This feature would potentially increase its application to the other type of surgeries (e.g., ESinS, microscopic ear surgery, etc) after appropriate and dedicated validation studies.

Conclusion

MBS is an effective method to rate surgical field during EEarS. It was proved to be easily and readily understandable, with good-to-excellent intra-rater and inter-rater reliability. It does not depend on the use of a specific instrument (i.e., suction) for the rating. Hence, although validated for EEarS, its application could potentially extend to other surgical fields.

Funding None.

Compliance with ethical standards

Conflict of interest All the authors have nothing to disclose.

References

1. Qiao H, Chen J, Li W et al (2016) Intranasal atomised dexmedetomidine optimises surgical field visualisation with decreased blood loss during endoscopic sinus surgery: a randomized study. Rhinology 54:38–44
2. Wormald PJ, van Renen G, Perks J et al (2005) The effect of the total intravenous anesthesia compared with inhalational anesthesia on the surgical field during endoscopic sinus surgery. Am J Rhinol 19:514–520
3. Halderman AA, Sindwani R, Woodard TD (2015) Hemorrhagic Complications of Endoscopic Sinus Surgery. Otolaryngol Clin North Am 48:783–793
4. Kelly EA, Gollapudy S, Riess ML, Woehlck HJ, Loehrl TA, Poetker DM (2013) Quality of surgical field during endoscopic sinus surgery: a systematic literature review of the effect of total intravenous compared to inhalational anesthesia. Int Forum Allergy Rhinol 3(6):474–481
5. Boezaart AP, van der Merwe J, Coetzee A (1995) Comparison of sodium nitroprusside- and esmolol-induced controlled hypotension for functional endoscopic sinus surgery. Can J Anaesth 42:373–376
6. Athanasiadis T, Beule A, Embate J, Steinmeier E, Field J, Wormald PJ (2008) Standardized video-endoscopy and surgical field grading scale for endoscopic sinus surgery: a multi-centre study. Laryngoscope 118(2):314–319

7. Bujang MA, Baharum N (2017) Guidelines of the minimum sample size requirements for Cohen's Kappa. EBPH. https://doi.org/10.2427/12267
8. Gisev N, Bell JS, Chen TF (2013) Interrater agreement and interrater reliability: key concepts, approaches, and applications. Res Soc Adm Pharm 9(3):330–338
9. Cichetti Dv (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 6(4):284–290
10. Cohen JA (1960) Coefficient of agreement for nominal scales. Educ Psychol Meas. https://doi.org/10.1177/001316446002000104
11. Sajedi P, Rahimian A, Khalili G (2016) Comparative evaluation between two methods of induced hypotension with infusion of remifentanil and labetalol during sinus endoscopy. J Res Pharm Pract 5(4):264–271
12. Govindaraj S, Adappa ND, Kennedy DW (2010) Endoscopic sinus surgery: evolution and tehcnical innovations. J Laryngol Otol 124(3):242–250
13. Beule AG, Wilhelmi F, Kühnel TS, Hansen E, Lackner KJ, Hosemann W (2007) Propofol versus sevoflurane: bleeding in endoscopic sinus surgery. Otolaryngol Head Neck Surg 136(1):45–50
14. Eberhart LH, Folz BJ, Wulf H, Geldner G (2003) Intravenous anesthesia provides optimal surgical conditions during microscopic and endoscopic sinus surgery. Laryngoscope 113(8):1369–1373
15. Ahn HJ, Chung SK, Dhong HJ, Kim HY, Ahn JH, Lee SM, Hahm TS, Kim JK (2007) Comparison of surgical conditions during propofol or sevoflurane anaesthesia for endoscopic sinus surgery. Br J Anesth 100(1):50–54
16. McGuire DB (1984) The measurement of clinical pain. Nurs Res 33(3):152–156