

## RESPIRATION AND THE AIRWAY

# Cormack–Lehane classification revisited

R. Krage, C. van Rijn, D. van Groeningen, S. A. Loer, L. A. Schwarte and P. Schober\*

Department of Anaesthesiology, VU University Medical Center, De Boelelaan 1117, 1007 MB Amsterdam, The Netherlands

\* Corresponding author. E-mail: p.schober@vumc.nl

### Key points

- The Cormack–Lehane (CL) classification is frequently used to describe laryngeal view on direct laryngoscopy.
- The present study shows that many anaesthetists lacked relevant knowledge.
- The reproducibility of CL classification was limited, with a poor intra-observer reliability and a fair inter-observer reliability.
- These results question the validity of using this classification in clinical and research settings.

**Background.** The Cormack–Lehane (CL) classification is broadly used to describe laryngeal view during direct laryngoscopy. This classification, however, has been validated by only a few studies reporting inconclusive data concerning its reliability. This discrepancy between widespread use and limited evidence prompted us to investigate the knowledge about the classification among anaesthesiologists and its intra- and inter-observer reliability.

**Methods.** One hundred and twenty interviews were performed at a major European anaesthesia congress. Participants were interviewed about their general knowledge on grading systems to classify laryngeal view during laryngoscopy and were subsequently asked to define the grades of the CL classification. Inter- and intra-observer reliabilities were tested in 20 anaesthesiologists well familiar with the CL classification, who performed 100 laryngoscopies in a full-scale patient simulator.

**Results.** Although 89% of interviewed subjects claimed to know a classification to describe laryngeal view during laryngoscopy, 53% were able to name a classification. When specifically asked about the CL classification, 74% of the interviewed subjects stated to know this classification, whereas 25% could define all four grades correctly. In the simulator-based part of the study, inter-observer reliability was fair with a  $\kappa$  coefficient of 0.35 and intra-observer reliability was poor with a  $\kappa$  of 0.15.

**Conclusions.** The CL classification is poorly known in detail among anaesthesiologists and reproducibility even in subjects well familiar with this classification is limited.

**Keywords:** airway; larynx, laryngoscopy; larynx, vocal cords

Accepted for publication: 25 March 2010

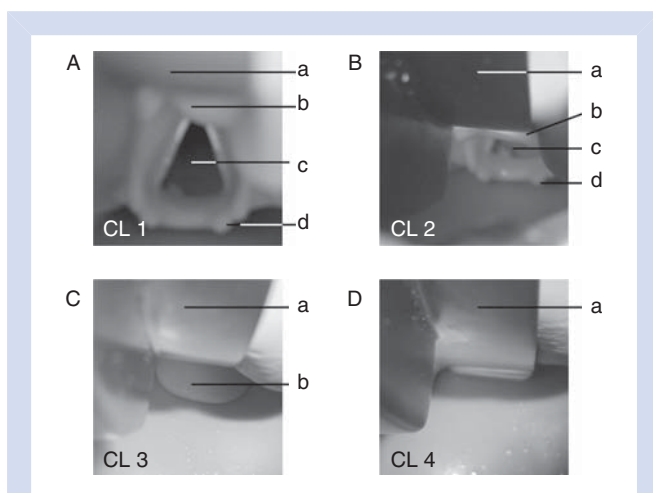
### Introduction

Tracheal intubation, most commonly performed using a direct laryngoscopy technique, is the gold standard in securing the airway and is considered mandatory in a variety of patient populations and operations. In this context, visibility of the glottis is often documented to describe intubating conditions. The Cormack–Lehane (CL) classification (Fig. 1) is a grading system commonly used to describe laryngeal view during direct laryngoscopy.<sup>1</sup> First published in 1984, it has since then become the gold standard for airway classification in clinical practice and in airway-related research.<sup>2–6</sup> However, despite its widespread use, the CL classification has not been fully validated. The few existing studies report inconclusive data in terms of inter- and intra-observer reliabilities.<sup>7–10</sup> One reason could be that anaesthesiologists are insufficiently familiar with the four different grades to allow correct practice, and modifications and contradictory definitions and figures in the literature may contribute to incorrect application. We therefore tested the hypothesis that the correct definition of the grades used in the CL

classification is poorly known among anaesthesiologists. In a second step, we aimed to test the inter- and intra-observer reliabilities of the CL classification in a group of anaesthesiologists well familiar with the correct definitions of all four grades.

### Methods

To test our hypothesis that the CL classification is poorly known in detail among anaesthesiologists, 120 interviews were performed at the Annual Meeting of the European Society of Anaesthesiology (ESA) 2008 in Copenhagen, Denmark. A sample size of 96 participants was computed to reach a confidence interval of 10% on a 95% confidence level (The Survey System 9.5, Creative Research Systems, Petaluma, CA, USA). The ESA congress was attended by more than 5000 delegates from all over the world. Participants were randomly asked to answer questions on a voluntary and anonymous basis. The questionnaires consisted of eight items, which were categorical and open answer questions (Table 1). These questions had previously been tested



**Fig 1** CL grades 1 (A), 2 (B), 3 (C), and 4 (D) in the SimMan™ human patient simulator. a, laryngoscope blade; b, epiglottis; c, glottic opening; d, arytenoid cartilages. According to the original definition by Cormack and Lehane, most of the glottic opening can be seen with grade 1. In grade 2, only the posterior portion of the glottis or only arytenoid cartilages are visible. In grade 3, only the epiglottis but no portion of the glottis is visible, whereas in grade 4, neither the glottis nor the epiglottis can be seen.

and re-adjusted in an internal audit process to optimize explicitness and conciseness. First, we asked whether any classifications were known at all to describe visibility of the glottis or laryngeal structures during laryngoscopy. Subsequently, participants were specifically interviewed about the CL classification and were requested to define all four grades. The last part of the questionnaire consisted of demographic items.

Inter- and intra-observer reliabilities of the CL classification were tested with volunteers performing laryngoscopies on the SimMan™ full-scale patient simulator (Laerdal Medical Corporation, Stavanger, Norway). Several airway-related features of the simulator such as mouth opening, neck mobility, tongue volume, and pharyngeal obstruction can be changed, enabling the SimMan™ to reproducibly simulate normal intubating conditions and a wide range of difficult airways.<sup>11</sup> One staff anaesthesiologist and one nurse anaesthetist with ample laryngoscopy experience tested a variety of settings to obtain typical CL grade 1, 2, 3, and 4 views (Fig. 1) during direct laryngoscopy with a Macintosh blade size #3 attached to an adult size laryngoscope handle (Heine Optotechnik, Herrsching, Germany). The four settings, which were found to match to CL grades 1 to 4, were re-evaluated and confirmed by two other staff anaesthesiologists.

We recruited 20 physicians from the Department of Anaesthesiology of the VU University Medical Centre in Amsterdam, the Netherlands, to participate on a voluntary basis and with informed consent. All of the participants regularly perform tracheal intubations in the operating theatre and are generally familiar with the CL classification. To

recapitulate the classification, the participants received a standardized briefing in which the definition of all grades was explained and a figure of the different grades was presented. Subsequently, each participant was asked to perform five laryngoscopies. In random order, CL grades 1, 3, and 4 were presented once, whereas grade 2 was presented twice. Randomization was performed by drawing five labelled cards from an opaque envelope and the order in which the cards were drawn determined the order of presented CL grades. The same manikin was used for all participants, which was in the supine position on an operating table with the head in the sniffing position (Fig. 2). All participants used the same laryngoscope handle and blade, and adequate lighting of the blade was ensured at all times. Participants were allowed to look without time limit; however, external laryngeal manipulation was not allowed. After each laryngoscopy, the participant announced the observed CL grade, which was then documented by the research team.

### Statistical Analysis

Results were analysed by the Prism 4.0 statistical package (GraphPad Software, San Diego, CA, USA) and two statistical calculators (QuickCalcs, GraphPad Software and Online Kappa Calculator, Randolph, J. J. 2008, retrieved October 27, 2009, from <http://justus.randolph.name/kappa>). Categorical data were compared using Fisher's exact test. A two-sided *P*-value of <0.05 was considered statistically significant.

The agreement between the preset CL grade in the simulator as defined by the investigator group and the grade actually seen by the participant was assessed by Cohen's  $\kappa$  coefficient for two raters.<sup>12</sup> Intra-rater agreement was defined as the agreement within each subject between his or her first and second assessments of preset CL grade 2. Inter-rater agreement between all 20 participants was assessed with a multirater variant of  $\kappa$  as described by Siegel and Castellan.<sup>13</sup>

### Results

Interviews were performed with 120 participants. Three participants without anaesthesiological background (two medical students and one internist) were excluded, leaving 117 participants [57 females, 60 males, age 42 (24–65) yr] eligible for data analysis. These 117 participants were 93 anaesthesiology specialists and 24 anaesthesiology residents from 39 nations. Three nations were highly represented accounting for about one-third of all participants: Great Britain ( $n=14$ ), Germany ( $n=14$ ), and the Netherlands ( $n=13$ ). Forty-nine participants were from the remaining 24 countries of the European Union (EU) and 27 were from non-EU countries and regions, including the USA, Canada, Latin America, Asia, Australia, and New Zealand.

One hundred and four participants (89%) claimed to know a classification to describe the visibility of laryngeal structures during laryngoscopy; however, only 62 (53%) were actively able to name a correct classification. In all 62 cases, the CL

**Table 1** Items of the anonymous questionnaire used for the survey performed at the Annual Meeting of the European Society of Anaesthesiology 2008. The questionnaire consisted of three pages. Participants were not allowed to return to a previous page to modify their initial answer

Page #	Item #	Question text
1	1	Do you know any classifications to describe the visibility of the larynx/glottis during laryngoscopy? Yes <input type="checkbox"/> No <input type="checkbox"/>
	2	If yes, please specify: _____
2	3	Do you know the Cormack–Lehane classification? Yes <input type="checkbox"/> No <input type="checkbox"/>
	4	If you know the Cormack–Lehane classification, could you please define the different grades? (free text)
3	5	Age: _____
	6	Gender: male <input type="checkbox"/> female <input type="checkbox"/>
	7	Country of origin: _____
	8	I am: Anaesthesiologist (medical specialist): <input type="checkbox"/> In training for anaesthesiology in my _____year of training <input type="checkbox"/> Other _____ <input type="checkbox"/>



**Fig 2** SimMan™ full-scale patient simulator positioning during laryngoscopies.

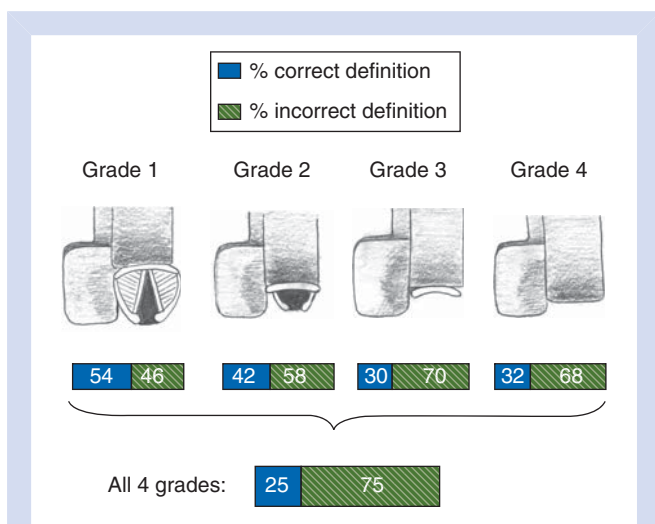
classification was named, whereas other existing classifications (e.g. the percentage of glottic opening score)<sup>7</sup> were not mentioned at all. The Mallampati classification,<sup>14</sup> which

is, since its modification by Samsoon and Young,<sup>15</sup> also a four-grade score based on the visibility of structures in the oral cavity including uvula and soft palate, was named by 29 participants (25%) and was the most frequently given incorrect answer. When specifically asked about the CL classification, 87 (74%) of the interviewed subjects claimed to know this classification. However, although 54% were able to define grade 1 correctly, only 42%, 30%, and 32% of the participants could correctly define grades 2, 3, and 4, respectively (Fig. 3). Merely 25% of all participants could define all four grades correctly, and even of those 87 participants who had claimed to know the CL classification, only 29 (33%) defined all grades correctly.

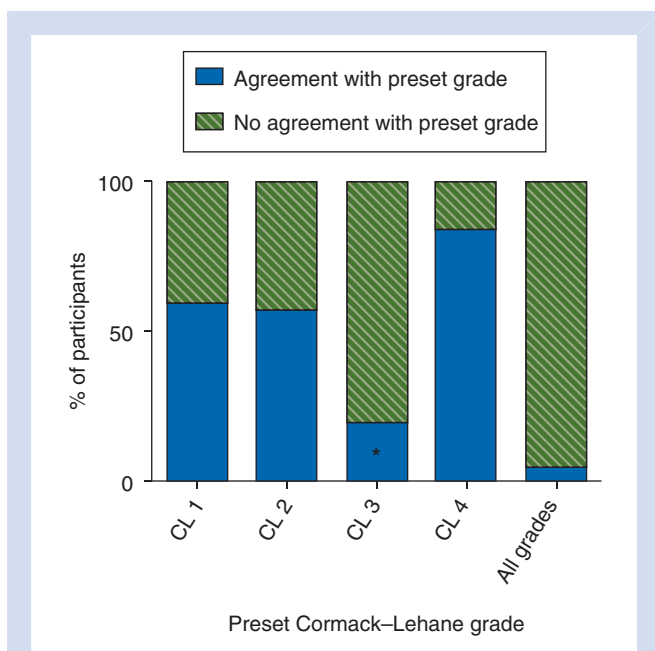
Knowledge about the CL classification did not differ between anaesthesiology specialists and residents ( $P > 0.05$ ) and did not differ between participants from the EU and non-EU countries ( $P > 0.05$ ).

### Simulator study

The 20 participants [five females, 15 males, age 36 (28–56) yr] were five anaesthesiology specialists and 15 anaesthesiology

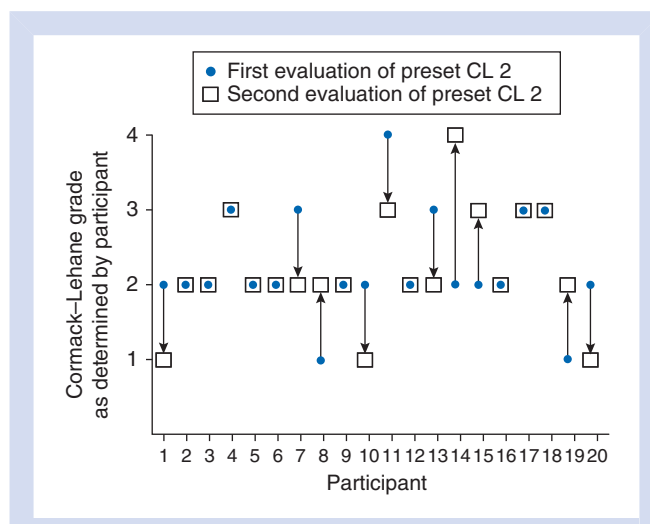


**Fig 3** Percentage of survey participants ( $n=117$ ) who defined CL grades 1, 2, 3, and 4 and all four grades correctly (blue portion of the bars) or incorrectly (green-striped portion of the bars).



**Fig 4** Agreement between preset CL grade and the grade observed by participants ( $n=20$ ). The blue portion of the columns represents the percentage of participants who showed agreement with the preset grade, whereas the green-striped portion represents the percentage of participants who showed no agreement. \* $P<0.05$  vs CL 1, CL 2, and CL 4.

residents. Each participant successfully completed five laryngoscopies according to the experimental protocol, resulting in a total of 100 laryngoscopies for data analysis. In 56 of the laryngoscopies, the observed grade conformed to preset grade, whereas no agreement was observed in 44 cases. Cohen’s  $\kappa$  coefficient of the overall agreement between the preset and the seen grade was 0.40 (95% confidence interval:



**Fig 5** Agreement of participants ( $n=20$ ) between their first (dot) and second (box) evaluation of preset CL grade 2. Ten of the 20 participants agreed in their first and second assessment, six participants assigned a lower grade ( $\downarrow$ ), and four participants a higher grade ( $\uparrow$ ) at the second evaluation.

0.27–0.53), indicating fair agreement.<sup>16</sup> Preset CL grade 1 was classified as grade 1 by 60% of the participants, grade 2 as grade 2 by 58%, and grade 4 as grade 4 by 85% of participants (Fig. 4). Preset grade 3 was significantly less often identified correctly (only classified as grade 3 by 20% of participants) than any other grade ( $P<0.05$ ). Merely one of the participants classified all airway conditions conforming to the preset grades. Multirater  $\kappa$  for the overall agreement between all 20 participants was 0.35.

CL grade 2 was assigned twice to each participant. Ten of the participants (50%) showed an agreement between their first and second observations of the identical airway setting (Fig. 5), resulting in a Cohen’s  $\kappa$  coefficient of 0.15 for the intra-observer agreement (95% confidence interval:  $-0.219$  to 0.524), indicating poor agreement.<sup>16</sup> Seven of these 10 laryngoscopists observed CL grade 2 on both occasions and three participants observed grade 3 during the first and second laryngoscopies of preset grade 2. The remaining 10 participants did not conform in their first and second grading of the identical airway setting.

## Discussion

We aimed to investigate the knowledge about the CL classification among anaesthesiologists and its reliability in a simulated clinical setting. Our survey shows that about three out of four anaesthesiologists claimed to know the CL classification, whereas only one out of four was able to define all grades correctly. Intra- and inter-observer reliabilities of the CL classification were tested with a full-scale patient simulator. Overall, the reproducibility of the CL classification was found to be limited, with a poor intra-observer reliability and a fair inter-observer reliability.

The survey had been prepared and performed in accordance with previously published recommendations for survey design.<sup>17 18</sup> The questionnaire was repeatedly evaluated by the study group to improve clarity and brevity. Familiarity with the CL classification may likely vary in different regions or hospitals, depending on local practice and training. Hence, performing the survey in one hospital or at a local congress would limit the validity of the data to the particular region where they were collected. We therefore chose to conduct the survey at the largest international European anaesthesiology congress, the Annual Meeting of the European Society of Anaesthesiology.

For the simulator-based study, we tested inter- and intra-observer reliabilities of the CL classification in a population of anaesthesiologists and anaesthesiology trainees of our own university hospital who are well familiar with this classification. To ensure that all participants had a uniform and correct knowledge of all four grades at the time of the study, they received a standardized briefing just before performing the laryngoscopies, in which the definitions of all grades were reviewed and a schematic figure showing the different grades was presented.

The human patient simulator SimMan™ allows predefined, reproducible, and real-time alterations of intubation conditions by manipulation of head and neck mobility, mouth opening, pharyngeal obstruction, and tongue volume.<sup>11</sup> This simulator is broadly used for airway management training<sup>19 20</sup> and has repeatedly been used as an instrument for airway-related research.<sup>21-25</sup> Hesselheldt and colleagues evaluated the SimMan™ and found that most of the anatomical features relevant to tracheal intubation, including ease of laryngoscope insertion, space in the oropharynx, and lifting force necessary to obtain glottic view, have an excellent realistic character.<sup>11</sup> Nevertheless, intubating conditions may differ from human subjects, for example, due to differences in tissue consistency or due to altered gliding properties of the laryngoscope on dry plastic compared with moist tongue mucosa. However, addressing the reliability of the CL classification in human subjects would require multiple laryngoscopies to be performed in the same patient. Besides ethical concerns with this approach, intubating conditions can change during repetitive laryngoscopy attempts due to changes in depth of anaesthesia and relaxation, oropharyngeal secretions, and swelling or bleeding due to previous manipulation. In contrast, the simulator allows standardized study conditions by reproducibly presenting identical airway settings to different participants or to the same participant at different time points. Regardless of potential differences between simulator and patient airway characteristics, all participants faced the same intubating condition with any given airway setting.

The question arises whether the preset airway settings as defined by the investigator group accurately reflect the four CL grades. The settings derive from expert consensus within the research group as they were worked out by two experienced SimMan™ operators with ample laryngoscopy experience and were subsequently confirmed by two other

experienced staff anaesthesiologists (see Fig. 1 for photo-documentation). A better-than-preset view was sometimes principally possible (e.g. view of the epiglottis in preset grade 4), however, only if the laryngoscopy was performed with extensive leverage on the maxillary incisors. All participants of the study were proficient with direct laryngoscopy, used a state-of-the-art left-handed laryngoscopy technique, and did not cause any damage to the breakable teeth of the simulator. Nevertheless, different participants graded identical preset airway conditions differently and one half of the participants graded the identical preset CL grade 2 differently at the second occasion. This variability between participants and within participants, rather than the accuracy of the preset grade, determines the inter- and intra-observer agreements of the CL classification.

All participants performed five laryngoscopies. Increasing fatigue of the left arm or decreasing concentration during repetitive laryngoscopy attempts could account for deteriorating laryngoscopy performance and hence deteriorating visibility of laryngeal structures. This effect was counteracted by randomization of the order in which the preset grades were presented to participants. Moreover, of the 10 participants who did not conform in their first and second grading of preset grade 2, six assigned a lower grading and four a higher grading at the second occasion (Fig. 5). This observation does not support the concept that fatigue may have led to relevant deteriorations in laryngoscopy performance over time. However, concerns about possible effects of fatigue if participants would have to perform more laryngoscopies prevented us from showing all grades twice to each participant and limited our assessment of intra-observer reliability to CL grade 2.

When initially describing their classification, Cormack and Lehane did not intend to propose a general classification for anaesthetists to document intubating conditions. Rather, it was intended for beginners to facilitate airway training and to improve intubating performance in cases of difficult tracheal intubation in obstetrics.<sup>1 26</sup> Nevertheless, this classification has inadvertently found its way into leading textbooks of airway management, anaesthesiology, intensive care, and emergency medicine as a classification to document intubating conditions.<sup>2-6</sup> For this purpose, however, the classification has only poorly been validated. Its validity has often intuitively been accepted, despite the lack of evidence. The CL classification is not only commonly used in clinical practice, but is also frequently used in airway-related research to describe and compare visibility of laryngeal structures.<sup>27-36</sup> This suggests a broad acceptance of this classification not only among clinicians but also among authors, reviewers, and editors of scientific journals. The obvious discrepancy between the widespread use of the CL classification and limited evidence to support its use prompted us to investigate the knowledge about the classification among anaesthesiologists and its intra- and inter-observer reliabilities.

One reason for poor knowledge of the CL classification among the participants in our study could be that this

classification is commonly mistaken with other similar classifications, such as the Mallampati classification, which was incorrectly named by one-fourth of our participants. Moreover, a certain ambiguity between CL grades 1 and 2 and several modifications of the CL classification, such as dividing grade 2 and/or grade 3 into grade 2a/2b and 3a/3b, respectively,<sup>37 38</sup> and contradicting definitions and figures in literature, may also likely contribute to confusion among anaesthesiologists. Survey results were similar between anaesthesiologists from different regions of the world and between medical specialists and doctors in training. Although we are aware that our survey was not powered to detect statistically significant differences between these populations, this may suggest that the rather poor knowledge about the CL classification we observed is not limited to certain regions or to certain training levels, but that it may rather apply to physicians from all parts of the world and to specialists and trainees alike. In a similar survey performed by Cohen and colleagues,<sup>39</sup> especially focused on the definition of grades 2 and 3, only 43% of 120 interviewed anaesthesiologists could correctly name the difference between grades 2 and 3 and that 45% defined a picture with a grade 2 view correctly.<sup>39</sup> Our results show a similarly limited knowledge among anaesthesiologists 15 yr after the study by Cohen and colleagues, suggesting that their conclusions still hold true today, despite an ongoing widespread use of the grading system in the meantime.

Inter- and intra-observer reliabilities of the CL classification were assessed by  $\kappa$  coefficients as a chance-corrected measure of agreement between categorical items, with a value of 1 indicating perfect agreement and 0 indicating no agreement better than chance. Values above 0.6 are generally accepted to indicate good agreement.<sup>16</sup>  $\kappa$  was 0.40 for the agreement between preset and observed CL grade and 0.35 for the agreement between the 20 participants, which both can be interpreted as a fair inter-observer agreement.<sup>16</sup> Interestingly, participants significantly less often correctly identified preset grade 3 than any other grade, and grade 3 was also the grade most often defined incorrectly in the survey. This may suggest that especially grade 3 imposes difficulties on users of the CL classification. Intra-observer agreement was poor with a  $\kappa$  of 0.15.

Four studies have previously addressed the reliability of the CL classification.<sup>7–10</sup> Levitan and colleagues<sup>7</sup> instructed five emergency physicians to rate 50 slides showing laryngeal views according to CL grades 1, 2, and 3. These 50 slides consisted of 25 different images and a duplicate of each image. For the purpose of the study, the CL classification was slightly modified and grade 1 was specifically defined as a view of the full glottic opening including the anterior commissure of the vocal cords.  $\kappa$  was 0.71 for intra-observer reproducibility and 0.59 for inter-observer reliability. The omission of CL grade 4 and the explicit distinction between grades 1 and 2, which is unambiguous in contrast to the original definition by Cormack and Lehane, may in part explain the rather high  $\kappa$  values in this study. The study by Ochroch and colleagues<sup>8</sup> used the same 25 pairs

of slides, which were presented to seven anaesthesiologists. As in the previous study, CL grade 4 was omitted; however, the original CL grading was used to define grades 1 and 2. An excellent intra-observer agreement was found ( $\kappa$  0.83); however, inter-observer reliability was poor ( $\kappa$  0.16).

The study by O'Shea and colleagues<sup>9</sup> assessed the reliability of the CL classification among seven paramedics. The experimental protocol closely parallels the methods applied by the first two studies, and the same 25 pairs of slides were used for this study. In contrast to the earlier two studies, all four grades of the CL classification were used. However, the authors again modified the CL score and defined grade 1 as a view of the full glottic opening including the anterior commissure of the vocal cords. Intra-rater reliability was assessed by calculating individual  $\kappa$  values for each participant, which ranged from 0.37 and 0.90; however, an overall or average  $\kappa$  value is not reported. Inter-rater reliability was poor with a  $\kappa$  value of 0.22.

The fourth study by George and colleagues<sup>10</sup> used a different approach. In patients undergoing elective surgery, tracheal intubation was recorded by a head-mounted camera that aligns with the laryngoscopist's line of sight. A video material from 26 patients was duplicated and the videos and duplicates were presented to the original laryngoscopist and to an independent anaesthesiologist. CL grades scored at initial laryngoscopy were compared with grades scored by the original laryngoscopist and the independent anaesthesiologist during review of the video material. From these multiple comparisons, three  $\kappa$  values were calculated for the intra-observer agreement (0.63, 0.64, and 0.74) and two  $\kappa$  values were calculated for the inter-rater agreement (0.70 and 0.61). However, in 21 out of 26 assessed cases, the initial grading during laryngoscopy and all four assessments of the video material from these laryngoscopies were consistently scored as CL grade 1, suggesting that the visibility of glottic structures was obviously very good in the majority of cases. In contrast, grade 4 was not scored at all and grade 3 was only scored in two patients in this study. This may suggest that the rather good reproducibility could mainly be due to the high number of patients in whom most of the glottic opening was visible during laryngoscopy and that this finding cannot necessarily be extrapolated to patients with airways corresponding to CL grades 3 and 4.

Although the range of  $\kappa$  values reported in these four studies is broad and therefore rather inconclusive, three of the four studies repeatedly reported  $\kappa$  values above 0.6. In contrast, our data demonstrate a limited inter- and intra-observer reproducibility of the CL classification. This discrepancy may in part be due to overestimations of  $\kappa$  values in earlier studies due to modifications of the definition of grade 1, due to omission of grade 4, or due to a marked overrepresentation of grade 1 as described above. However, the most striking difference between our study and previous studies is that our participants had to perform laryngoscopy themselves, resembling the actual clinical situation in which the classification is used. In contrast, previous studies used photo and video material, disregarding the influence of

stereoscopic vision and laryngoscopy itself on glottic visibility. It is obvious that not only the anatomy of the patient, but also the laryngoscopy technique has a major influence on glottic visibility. A variety of factors such as used force, skills, experience, and training effects contribute to the laryngoscopy technique, and individual differences in these factors may likely account for some of the inter-observer variability in our participants and in clinical practice alike. In this context, Williams and colleagues<sup>40</sup> pointed out that the incidence of CL grades 3 and 4 varies from 0.3% to 13% in several studies and conclude that this difference is due to differences in the laryngoscopic technique because the studied populations were comparable. Similarly, intra-individual differences in performing two laryngoscopies on preset grade 2 in our study may have led to a different visibility of the larynx on both occasions and may in part also account for the observed intra-observer variability.

The uncertainty among anaesthesiologists about the CL classification observed in the survey and its limited reliability under standardized conditions resembling the clinical situation restrict its clinical usefulness. A simpler classification may be better suitable for routine clinical practice, and such a simple grading system has been proposed by Cormack:<sup>26</sup> 'green light' cases can have various degrees of difficulty but should not cause failed intubation with an adequately trained anaesthesiologist. 'Red light' cases on the other hand did impose major difficulties and can cause failure during subsequent intubations. In our institution, we are successfully using a similar two-grade system since many years and additionally document whether mask ventilation was easily possible, because this may be much more relevant to ensure adequate oxygenation during induction than any other factor. After difficult intubations, which Cormack would grade a 'red light' case, we routinely document all relevant information as free text, for example, what actually caused the difficulty (subglottic stenosis, for example, may cause difficult intubation even if the glottis can be entirely seen), how many intubation attempts were necessary, how many anaesthetists were involved, what was their level of training, and which technique or intubation aids were finally successful to accomplish tracheal intubation. Although there is a tendency in medicine to use all sorts of classifications to summarize and condense complex circumstances into simple scores, we are convinced that free-text documentation will be more informative for subsequent anaesthesiologists than any single score derived from whatsoever classification.

Although free-text documentation may be more appropriate for the single case, a reliable and reproducible classification facilitates registration and comparison of intubating conditions across cases for epidemiological purposes. Moreover, such a classification could allow objective reporting of the laryngeal view or intubating conditions in airway-related research. Currently, the CL classification is broadly used for this purpose. Our data, however, suggest that the usefulness of this classification to objectively report and compare intubating conditions in airway-related research is rather

limited and that CL grades reported in medical literature should be interpreted with care.

In conclusion, despite a widespread use of the CL classification, only a minority of anaesthesiologists demonstrated adequate knowledge on this grading system. The CL classification showed a fair inter-observer reliability and a poor intra-observer reliability, even when applied by physicians well familiar with this rating system under standardized conditions. These results question the validity of the CL classification to document laryngeal view during direct laryngoscopy.

## Conflict of interest

None declared.

## Funding

This work was supported by departmental funding only.

## References

- 1 Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia* 1984; **39**: 1105–11
- 2 Gal TJ. Airway management. In: Miller DM, Fleisher LA, Johns RA, Savarese JJ, Winer-Kronish JP, Young WL, eds. *Miller's Anesthesia*. Philadelphia, PA: Elsevier, 2005; 1617–52
- 3 Benumof JL. Definition and incidence of the difficult airway. In: Benumof JL, ed. *Airway Management*. St Louis, MO: Mosby, 1996; 121–5
- 4 Rosenblatt WH. Airway management. In: Barash PG, Cullen BF, Stoelting RK eds. *Clinical Anesthesia*. Philadelphia, PA: Lippincott Williams & Wilkins, 2006; 595–642
- 5 Walls RM. Airway. In: Marx JA, Hockberger RS, Walls RM, eds. *Rosen's Emergency Medicine: Concepts and Clinical Practice*. Philadelphia, PA: Mosby Elsevier, 2010; 3–22
- 6 Kauer S, Heard SO. Airway management and endotracheal intubation. In: Irwin RS, Rippe JM, eds. *Irwin and Rippe's Intensive Care Medicine*. Philadelphia, PA: Wolters Kluwer, Lippincott Williams & Wilkins, 2008; 3–18
- 7 Levitan RM, Ochroch EA, Kush S, Shofer FS, Hollander JE. Assessment of airway visualization: validation of the percentage of glottic opening (POGO) scale. *Acad Emerg Med* 1998; **5**: 919–23
- 8 Ochroch EA, Hollander JE, Kush S, Shofer FS, Levitan RM. Assessment of laryngeal view: percentage of glottic opening score vs Cormack and Lehane grading. *Can J Anaesth* 1999; **46**: 987–90
- 9 O'Shea JK, Pinchall ME, Wang HE. Reliability of paramedic ratings of laryngoscopic views during endotracheal intubation. *Prehosp Emerg Care* 2005; **9**: 167–71
- 10 George RB, McKeen D, Law JA. Laryngoscopic evaluation with the Airway Cam. *Can J Anaesth* 2006; **53**: 512–5
- 11 Hessfeldt R, Kristensen MS, Rasmussen LS. Evaluation of the airway of the SimMan full-scale patient simulator. *Acta Anaesthesiol Scand* 2005; **49**: 1339–45
- 12 Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960; **20**: 37–46
- 13 Siegel S, Castellan NJ. Measures of association and their tests of significance. *Nonparametric Statistics for Behavioural Science*. Singapore: McGraw-Hill Publishing, 1988; 224–312
- 14 Mallampati SR, Gatt SP, Gugino LD et al. A clinical sign to predict difficult tracheal intubation: a prospective study. *Can Anaesth Soc J* 1985; **32**: 429–34

- 15 Samssoon GL, Young JR. Difficult tracheal intubation: a retrospective study. *Anaesthesia* 1987; **42**: 487–90
- 16 Altman DG. Some common problems in medical research. *Practical Statistics for Medical Research*. Boca Raton, FL: Chapman & Hall/CRC Press, 1991; 396–439
- 17 Burmeister LF. Principles of successful sample surveys. *Anesthesiology* 2003; **99**: 1251–2
- 18 Evans SJ. Good surveys guide. *Br Med J* 1991; **302**: 302–3
- 19 Grenvik A, Schaefer JJ, DeVita MA, Rogers P. New aspects on critical care medicine training. *Curr Opin Crit Care* 2004; **10**: 233–7
- 20 Timmermann A, Eich C, Nickel E et al. Simulation and airway management. *Anaesthesist* 2005; **54**: 582–7
- 21 Wackett A, Anderson K, Thode H. Bullard laryngoscopy by naive operators in the cervical spine immobilized patient. *J Emerg Med* 2005; **29**: 253–7
- 22 Maharaj CH, Ni CM, Higgins BD, Harte BH, Laffey JG. Tracheal intubation by inexperienced medical residents using the Airtraq and Macintosh laryngoscopes—a manikin study. *Am J Emerg Med* 2006; **24**: 769–74
- 23 Savoldelli GL, Schiffer E, Abegg C, Baeriswyl V, Clergue F, Waeber JL. Learning curves of the Glidescope, the McGrath and the Airtraq laryngoscopes: a manikin study. *Eur J Anaesthesiol* 2009; **26**: 554–8
- 24 Malik MA, Hassett P, Carney J, Higgins BD, Harte BH, Laffey JG. A comparison of the Glidescope, Pentax AWS, and Macintosh laryngoscopes when used by novice personnel: a manikin study. *Can J Anaesth* 2009; **56**: 802–11
- 25 Powell L, Andrzejowski J, Taylor R, Turnbull D. Comparison of the performance of four laryngoscopes in a high-fidelity simulator using normal and difficult airway. *Br J Anaesth* 2009; **103**: 755–60
- 26 Cormack RS. Laryngoscopy grades. *Anaesthesia* 1999; **54**: 911–2
- 27 Eich C, Roessler M, Nemeth M, Russo SG, Heuer JF, Timmermann A. Characteristics and outcome of prehospital paediatric tracheal intubation attended by anaesthesia-trained emergency physicians. *Resuscitation* 2009; **80**: 1371–7
- 28 Soyuncu S, Eken C, Cete Y, Bektas F, Akcimen M. Determination of difficult intubation in the ED. *Am J Emerg Med* 2009; **27**: 905–10
- 29 Malik MA, Subramaniam R, Maharaj CH, Harte BH, Laffey JG. Randomized controlled trial of the Pentax AWS, Glidescope, and Macintosh laryngoscopes in predicted difficult intubation. *Br J Anaesth* 2009; **103**: 761–8
- 30 Khan ZH, Mohammadi M, Rasouli MR, Farrokhnia F, Khan RH. The diagnostic value of the upper lip bite test combined with sternal distance, thyromental distance, and interincisor distance for prediction of easy laryngoscopy and intubation: a prospective study. *Anesth Analg* 2009; **109**: 822–4
- 31 Macnair D, Baraclough D, Wilson G, Bloch M, Engelhardt T. Pediatric airway management: comparing the Berci-Kaplan Video Laryngoscope with direct laryngoscopy. *Paediatr Anaesth* 2009; **19**: 577–80
- 32 Ali Z, Bithal PK, Prabhakar H, Rath GP, Dash HH. An assessment of the predictors of difficult intubation in patients with acromegaly. *J Clin Neurosci* 2009; **16**: 1043–5
- 33 Malin E, Montblanc J, Ynineb Y, Marret E, Bonnet F. Performance of the Airtraq laryngoscope after failed conventional tracheal intubation: a case series. *Acta Anaesthesiol Scand* 2009; **53**: 858–63
- 34 Asai T, Liu EH, Matsumoto S et al. Use of the Pentax-AWS in 293 patients with difficult airways. *Anesthesiology* 2009; **110**: 898–904
- 35 Huh J, Shin HY, Kim SH, Yoon TK, Kim DK. Diagnostic predictor of difficult laryngoscopy: the hyomental distance ratio. *Anesth Analg* 2009; **108**: 544–8
- 36 Lange M, Frommer M, Redel A et al. Comparison of the Glidescope and Airtraq optical laryngoscopes in patients undergoing direct microlaryngoscopy. *Anaesthesia* 2009; **64**: 323–8
- 37 Cook TM. A grading system for direct laryngoscopy. *Anaesthesia* 1999; **54**: 496–7
- 38 Yentis SM, Lee DJ. Evaluation of an improved scoring system for the grading of direct laryngoscopy. *Anaesthesia* 1998; **53**: 1041–4
- 39 Cohen AM, Fleming BG, Wace JR. Grading of direct laryngoscopy. A survey of current practice. *Anaesthesia* 1994; **49**: 522–5
- 40 Williams KN, Carli F, Cormack RS. Unexpected, difficult laryngoscopy: a prospective survey in routine general surgery. *Br J Anaesth* 1991; **66**: 38–44