

Description and comparison of the performance of the upper lip bite test, the ratio of height to thyromental distance and other methods of preoperative airway assessment in a Nigerian population – a pilot study

OF Dada, AF Faponle, AT Adenekan

Department of Anaesthesia, Obafemi Awolowo University Teaching Hospitals Complex, Ile-Ife, Nigeria

Corresponding author, email: funkedada0@yahoo.com

Background: The ability to reliably predict difficult intubation can prevent airway related adverse outcomes. The upper lip bite test (ULBT) and ratio of height to thyromental distance (RHTMD) are airway predictive tests that have been evaluated in Caucasian populations with promising results. This study determined the test performance of the ULBT and RHTMD and compared with commonly employed predictive tests (modified Mallampati [MMT], thyromental distance [TMD] and inter-incisor gap [IIG]) in adult Nigerians.

Methods: This was a cross-sectional study. Two hundred and sixteen consecutive, consenting ASA I-III patients presenting for surgery and planned for general anaesthesia with endotracheal intubation over a six month period in our teaching hospital, had all five airway assessment tests (ULBT, RHTMD, MMT, TMD and IIG) assessed preoperatively. The Cormack and Lehane grading was used to determine easy or difficult visualisation of the larynx (EVL or DVL). Sensitivity, specificity, predictive values and likelihood ratios of the tests were determined.

Results: The sensitivities, specificities, positive and negative predictive values and positive and negative likelihood ratios respectively for the tests were: ULBT (11.8%, 99.0%, 50.0%, 92.9%, 11.71 and 0.89), RHTMD (35.3%, 92.5%, 28.6%, 94.4%, 4.68 and 0.70), TMD (29.4%, 97.5%, 50.0%, 94.2%, 11.71 and 0.72), MMT (52.9%, 86.4%, 25.0%, 95.6%, 3.90 and 0.54), and IIG (11.8%, 97.0%, 25.0%, 92.8%, 3.90 and 0.91).

Conclusion: The modified Mallampati test had the highest sensitivity in this study, however all the tests evaluated showed only low to moderate sensitivity. The ULBT and RHTMD had low sensitivities in this population. However, based on the high positive likelihood ratios of the ULBT and the TMD, whenever positive, these tests do show a high probability of DVL and these two tests would probably perform similarly in this population. Anthropometric differences may account for differences in performance of preoperative airway assessment tests in various populations.

Keywords: difficult laryngoscopy, airway assessment, predictive airway screening tests

Introduction

Maintaining a patent airway and adequate gas exchange in patients undergoing anaesthesia remains a vital responsibility of every anaesthetist. Endotracheal intubation is a procedure regularly carried out by anaesthetists; nevertheless, difficulty sometimes occurs with glottic exposure and tracheal intubation.¹

Adequate preoperative airway assessment and appropriate planning may prevent unanticipated difficult airway and potentially catastrophic airway-related events. Closed claims analysis in healthy patients undergoing non-emergency surgeries revealed that 85% of airway-related events involved brain damage or brain death.² Thirty-three percent of deaths attributable solely to the process of anaesthesia have been related to the inability to safeguard the patency of the airway.³

The incidence of difficult intubation ranges between 1.5–13%.⁴⁻⁸ The American Society of Anesthesiologists' Task Force on difficult airway defined difficult laryngoscopy as when it is not possible to visualise any portion of the vocal cords during conventional laryngoscopy.¹ Difficult intubation was defined as a situation in which tracheal intubation requires multiple attempts in the presence or absence of tracheal pathology.¹

Recognition of the potentially fatal consequences of failed airway have led to the introduction of newer airway devices, algorithms, and research into the most accurate test(s) for preoperative airway evaluation.⁴⁻¹³ The adequacy of commonly used single screening tests (such as modified Mallampati, thyromental distance, inter-incisor gap, sternomental distance) has been questioned, likewise the applicability of complex tests like the Wilson score in day to day practice.^{14,15}

Merah et al.⁵ found that in West Africans, the best predictors of difficult laryngoscopy were the combination of modified Mallampati (MMT), thyromental distance (TMD) and inter-incisor gap (IIG) with the MMT being the most sensitive test. Two more recently introduced bedside screening tests – the upper lip bite test (ULBT) and ratio of height to thyromental distance (RHTMD) – have been studied in Caucasian populations.^{6,7,8} The ULBT assesses the ability of the lower incisors to bite the upper lip while evaluating for mandibular subluxation and buck teeth.⁶ The RHTMD was introduced, with the advantage over the thyromental distance, of allowing for individual body proportions.⁷ These two tests have been shown, in Caucasians, to be better predictors of difficult visualisation of the larynx than other single airway assessment tests, including the MMT, TMD and IIG.^{6,7,8}

Anthropological differences have been found to exist in various populations especially with respect to facial morphology. Hence, airway assessment tests useful in Caucasians may not always translate to a similar efficacy in African populations and vice versa. There appears to be a paucity of data on the validity of these two tests (ULBT and RHTMD) in the African population.

The primary objective of this study was therefore to describe the performance of the new screening tests (ULBT and RHTMD) in Nigerian adult population. The secondary objective was to compare the new screening tests with existing tests (MMT, TMD and IIG) in terms of reliable prediction of difficult visualisation of the larynx in this population. The authors hypothesised that the upper lip bite test and ratio of height to thyromental distance would perform differently as predictors of visualisation of the larynx than the commonly used tests (MMT, TMD and IIG) in the adult Nigerian population.

Materials and methods

This was a cross-sectional study conducted at Obafemi Awolowo University Teaching Hospitals Complex (OAUTHC), Ile-Ife. The hospital is a 687-bedded tertiary hospital. General surgery procedures, minimally invasive (laparoscopic) surgeries, obstetrics and gynaecological surgeries, neurosurgeries, cardiothoracic surgeries (including open heart surgeries), urological surgeries (including renal transplantation), orthopaedic surgeries (including joint replacement surgeries and spine surgeries), paediatric surgeries, plastic surgeries, ophthalmic surgeries, otorhinolaryngological surgeries and maxillofacial surgeries are performed at the hospital. Institutional ethics approval was obtained before commencement of the study. All eligible consecutive, consenting adult patients 18–65 years, American Society of Anesthesiologists (ASA) physical status I–III scheduled for elective general (endotracheal) anaesthesia were recruited over a six-month period (between March 1, 2015 and September 30, 2015).

Patients who refused to participate or could not carry out instructions, edentulous patients, patients who were unable to sit or stand erect, patients with gross anatomical airway malformations, obstetric patients, patients with contraindication to direct laryngoscopy or a history of difficult intubation, ASA IV or V patients, obese patients and patients with 'full stomach' were excluded from the study.

The study protocol was discussed extensively with four experienced anaesthetists covering the four main operating theatres for elective surgeries during the study period (who would be responsible for airway management) and the description of the Cormack and Lehane grading re-emphasised to them (including the use of visual aids) prior to commencement of the research. The principal investigator performed the preoperative airway assessment tests on all the patients but was not involved in the intubation of the patients.

Written informed consent was obtained from patients and information collected prospectively by means of a standardised

record sheet during the pre-anaesthetic visit by the investigator. All the patients had preoperative review (and performance of the airway assessment tests) carried out by the investigator. The upper lip bite test (ULBT) was graded with the patient sitting and head in neutral position.⁶ The patients were asked to bite the upper lip with the lower incisors as far upwards as possible. If the patient could bite the mucosa of the upper lip above the vermilion line, it was recorded as Grade I. If the patient could bite the upper lip with the lower incisors but below the vermilion line it was recorded as Grade II. If the patient could not bite the upper lip with the lower incisors, it was recorded as Grade III. Grade III was considered predictive of difficult visualisation of the larynx (DVL).

The ratio of height to thyromental distance (RHTMD) was measured as follows: the height of the patient was measured on a stadiometer. The thyromental distance was measured in centimetres with a tape measure. The height of the patient (in cm) was divided by the thyromental distance (in cm). A value greater than or equal to 25 was considered predictive of DVL.⁷

The modified Mallampati assessment was done with the patient sitting, head in neutral position, mouth fully open, and maximal tongue protrusion without phonating.¹⁶ Class III and IV were considered predictive of DVL. The thyromental distance was assessed by measuring, with a tape measure, the distance between the thyroid prominence and the mentum with the head fully extended and the mouth closed.¹⁷ If less than or equal to 6 cm, it was considered predictive of DVL.¹⁸ The inter-incisor gap was measured with the mouth of the patient fully open with the aid of calipers.¹⁹ Inter-incisor gap less than or equal to 4 cm was considered predictive of DVL.

All patients had baseline and intraoperative vital signs measured (pulse rate, blood pressure, respiratory rate, temperature, electrocardiography and Oxygen saturation) using the Dash 4000 multi-purpose monitor by General Electric Medical Systems (2003). Patients were pre-oxygenated and pre-medicated as per hospital protocol. Pressor response to laryngoscopy was attenuated with IV lidocaine 1.5 mg/kg. Induction of anaesthesia was done with IV propofol 2 mg/kg and endotracheal intubation facilitated with IV suxamethonium 1.5 mg/kg. After disappearance of fasciculations, laryngoscopy was done with the head in the sniffing position using a Macintosh size 3 or 4 blade by the anaesthetist. To ensure safety was not compromised, a fully equipped difficult airway cart was always available in each theatre and patients with overt features of difficulty (for instance gross anatomical airway malformations or any situation in which, per discussion with the attending anaesthetist before patient recruitment, they felt performance of an asleep induction or induction using intravenous anaesthesia would compromise safety). This was stated when discussing the study with the department before the commencement of the study but there were no such incidents during the course of the study as patients in this category were already embodied in the exclusion criteria and were therefore excluded from the study. The Cormack and Lehane²⁰ grading which was the gold standard to which the

pre-anaesthetic airway assessment tests were compared was recorded for all the patients. Grades 1 (full view of the glottis) and 2 (partial view of the glottis or arytenoids) were considered easy visualisation of the larynx (EVL) and Grades 3 (only the epiglottis visible) and 4 (neither the glottis nor epiglottis visible) considered difficult visualisation of the larynx (DVL). The best view obtained by the anaesthetist performing the laryngoscopy was what was recorded (after external laryngeal manipulations where necessary).

Following endotracheal intubation, confirmation of correct placement was by capnography. Anaesthesia proceeded normally as per hospital protocol with patients maintained on Isoflurane in oxygen/air, intravenous fluids as required, analgesia and muscle relaxation. At the end of the procedure, residual neuromuscular blockade was reversed in all patients with intravenous neostigmine 0.04 mg/kg, immediately preceded by intravenous atropine 0.02 mg/kg. Patients were extubated at the discretion of the attending anaesthetist and transferred to the post-anaesthesia care unit. Each patient was followed up until satisfactory recovery from anaesthesia.

Statistical analysis

Data was entered, screened, corrected and analysed using IBM Statistical Package for Scientific Solutions (SPSS) software version 20. Quantitative continuous variables were summarised using mean and standard deviation (mean \pm SD) while frequencies and percentages, n (%) were used for categorical variables. The sensitivity, specificity, positive predictive and negative predictive values of each of the airway assessment tests together with their 95% confidence intervals were calculated. Likelihood ratios were also determined and added as exploratory analysis, as well as Receiver Operating Characteristics (ROC) curve for the RHTMD even though these were not pre-specified.

The Pearson Chi-square test was used to detect any association between categorical patient's characteristics and airway visualisation while the independent samples 't' test was used to compare means of quantitative continuous variables between patients with easy visualisation and those with difficult visualisation. The level of statistical significance was determined at p-values less than 0.05.

Results

A total of 216 consecutive eligible patients within the study period were recruited for the study (see Figure 1). There were 57 males and 159 females. The incidence of DVL was 7.9% (17 of the 216 patients). There was a statistically significant difference in age between the difficult visualisation of the larynx (DVL) group (mean age 52.9 ± 11.1) and the easy visualisation of the

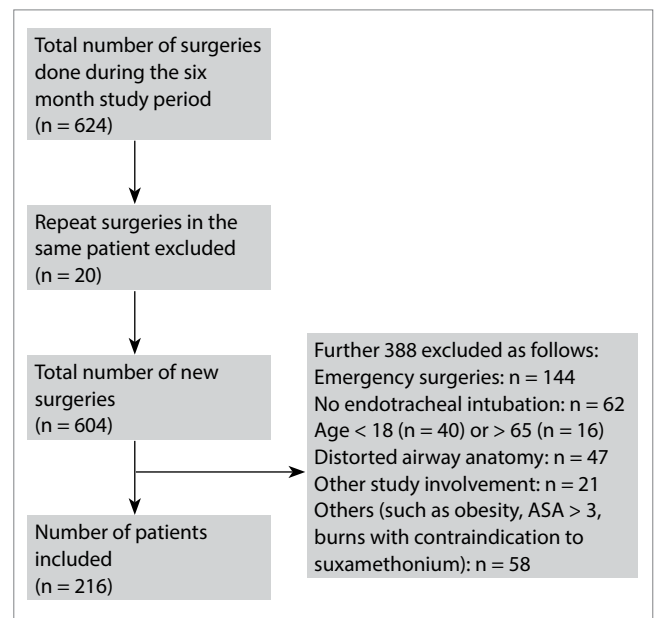


Figure 1. Prisma diagram

Table I. Demographic characteristics

Variables	DVL (n = 17)	EVL (n = 199)	t	P value
Age in years	52.9 \pm 11.1	40.4 \pm 13.1	3.822	< 0.001
Gender n (%)				
Male	6 (10.5)	51 (89.5)	0.753	0.385
Female	11 (6.9)	148 (93.1)		
Surgical type n (%)				
General Surgery	6 (6.0)	94 (94.0)	13.660	0.018
Gynaecological	3 (3.9)	73 (96.1)		
Head and Neck	3 (30.0)	7 (70.0)		
Orthopaedic/Plastic	3 (15.0)	17 (85.0)		
Urology	2 (25.0)	6 (75.0)		
Cardio-thoracic	0 (0.0)	2 (100.0)		
Weight (Kg)	65.0 \pm 9.7	65.1 \pm 8.6	-0.063	0.950
Height (m)	1.7 \pm 0.1	1.7 \pm 0.1	-0.780	0.436
Body mass index (Kg/m²)	23.8 \pm 2.9	23.5 \pm 3.2	0.387	0.699

DVL - Difficult Visualisation of the Larynx, EVL - Easy Visualisation of the Larynx

Table II. Individual predictors of DVL and EVL

Predictive test	DVL (n = 17)	EVL (n = 199)
ULBT		
Positive	2	2
Negative	15	197
RHTMD		
Positive	6	15
Negative	11	184
TMD		
Positive	5	5
Negative	12	194
MMT		
Positive	9	27
Negative	8	172
IIG		
Positive	2	6
Negative	15	193

DVL – Difficult Visualisation of the Larynx, EVL – Easy Visualisation of the Larynx, ULBT – Upper lip bite test, RHTMD – Ratio of Height to Thyromental Distance, TMD – Thyromental Distance, MMT – Modified Mallampati Test, IIG – Inter-Incisor Gap

larynx (EVL) group (mean age 40.4 ± 13.1) (P < 0.001). There was no statistically significant difference in gender, weight, height or body mass index between the DVL and EVL groups (Table I).

The ULBT correctly identified only two of the 17 patients with DVL in this study, likewise the IIG. The RHTMD identified six, MMT identified nine and TMD five of the patients with DVL. Majority of the tests correctly identified most of the patients with EVL (Table II).

Table III shows the sensitivities, specificities, positive and negative predictive values as well as positive and negative likelihood ratios for the five tests. The sensitivities, specificities, positive and negative predictive values and positive and negative likelihood ratios respectively for the tests were: ULBT (11.8%, 99.0%, 50.0%, 92.9%, 11.71 and 0.89), RHTMD (35.3%, 92.5%, 28.6%, 94.4%, 4.68 and 0.70), TMD (29.4%, 97.5%, 50.0%, 94.2%, 11.71 and 0.72) MMT (52.9%, 86.4%, 25.0%, 95.6%, 3.90 and 0.54) and IIG (11.8%, 97.0%, 25.0%, 92.8%, 3.90 and 0.91). The highest sensitivity was observed with the modified Mallampati test. The upper lip bite test and inter incisor gap had the least sensitivity. The highest

Table III. Sensitivity, specificity and predictive values of the airway predictors

Test	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-
ULBT	11.8 (3.3–34.3)	99.0 (96.4–99.7)	50.0 (15.0–85.0)	92.9 (88.7–95.7)	11.71 (1.76–77.98)	0.89 (0.74–1.06)
RHTMD	35.3 (17.3–58.7)	92.5 (87.9–95.4)	28.6 (13.8–50.0)	94.4 (90.2–96.8)	4.68 (2.09–10.49)	0.70 (0.49–1.00)
TMD	29.4 (13.3–53.1)	97.5 (94.3–98.9)	50.0 (23.7–76.3)	94.2 (90.1–96.6)	11.71 (3.76–36.47)	0.72 (0.53–0.98)
MMT	52.9 (31.0–73.8)	86.4 (81.0–90.5)	25.0 (13.8–41.1)	95.6 (91.5–97.7)	3.90 (2.21–6.89)	0.54 (0.33–0.90)
IIG	11.8 (3.3–34.3)	97.0 (93.6–98.6)	25.0 (7.1–59.1)	92.8 (88.4–95.6)	3.90 (0.85–17.87)	0.91 (0.76–1.08)

ULBT – Upper lip bite test, RHTMD – Ratio of Height to Thyromental Distance, TMD – Thyromental Distance, MMT – Modified Mallampati Test, IIG – Inter-Incisor Gap, PPV – Positive Predictive Value, NPV – Negative Predictive Value, LR – Likelihood Ratio

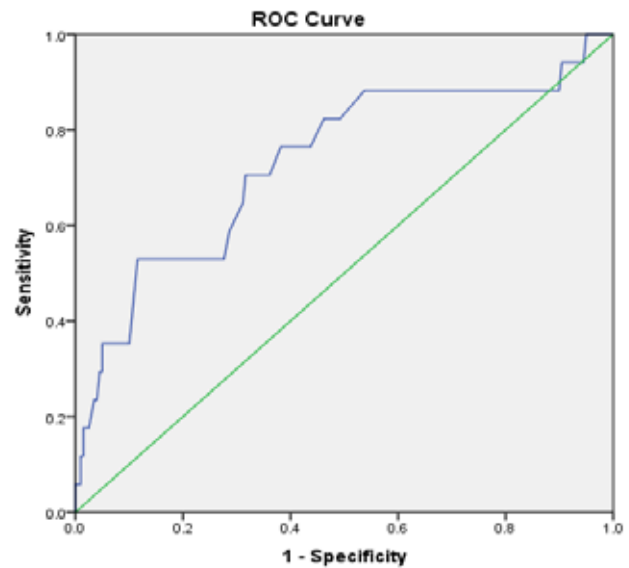


Figure 2. Receiver operating characteristics curve for the RHTMD; Area under the curve = 0.732

positive likelihood ratios were observed for the ULBT and TMD (11.71).

From the data obtained from this study, to determine the optimal cut-off point for the RHTMD that will detect difficult visualisation of the larynx in this population, the Receiver operating characteristics (ROC) curve was plotted and coordinates of the ROC Curve analysis obtained (Figure 1, Supplementary Table I). The area under the curve is 0.732 (confidence interval 0.592–0.872), indicating a fair separation of the population studied by the RHTMD into DVL and EVL. The optimal RHTMD for this population appears to be at a cut-off point of 24.7 (sensitivity 35%, false positives 5%) (Supplementary Table I).

Discussion

In the studied Nigerian population, the upper lip bite test (ULBT) and ratio of height to thyromental distance (RHTMD) had low sensitivities. The modified Mallampati test (MMT) had the highest sensitivity in this study. Most of the tests evaluated in this study had low to moderate sensitivities and low positive predictive values (PPV) albeit with high specificities and negative predictive values (NPV).

The ideal preoperative airway assessment test would be highly sensitive to predict majority of patients with DVL as well as

highly specific to be able to correctly identify patients with easy visualisation of the larynx (EVL). It should also have a high positive predictive value so that only a few patients with EVL are subjected to difficult intubation drills and a high negative predictive value (with few false negative predictions). The positive likelihood ratio (LR+) shows the strength of the probability of a test to identify DVL when present and the negative likelihood ratio (LR-) shows the strength of the probability of a test to show if DVL is absent.²¹ The likelihood ratios are less affected by prevalence of DVL than the positive and negative predictive values.

The overall incidence of DVL in this study was 7.9%. This is higher than the incidence reported by Merah et al.⁵ in a West African population but comparable with the incidence quoted in some other studies.^{7,22} Variations in the incidence of DVL have been attributed to different factors, such as different anthropometric features among populations, lack of uniformity in describing or grading laryngeal views, cricoid pressure application, head position, degree of muscle relaxation, and type or size of laryngoscope blade.^{22,23} In this study the Cormack and Lehane grading recorded was the best view obtained (following backward upward rightward pressure where necessary), while some other studies including the study by Merah et al.⁵ reported the laryngeal views before any external laryngeal manipulations. Variations such as these and others as stated above may account for the differences in incidence of DVL reported.

The upper lip bite test had a low sensitivity of 11.8%. It correctly identified majority (99%) of patients with EVL. However, because safety in airway management is more directly improved by ability to correctly predict difficult airway, this high specificity would be of secondary importance. Khan et al.⁶ originally described the ULBT as a simple, easily reproducible test with a sensitivity of 76.5% and specificity of 88.7%. This is different from the very low sensitivity recorded in the index study. The risk of inter-observer bias is thought to be more unlikely for the ULBT than most other airway screening tests due to its very definite demarcations for the grades. The low sensitivity of the ULBT in this present study is also different from the high sensitivity (74.63%) reported by Shah et al.⁸ in India. Eberhart et al.²⁴, on the other hand, in their study comparing the ULBT and the MMT on 1 106 patients in Germany found the sensitivity, specificity, PPV, and NPV obtained for the upper lip bite test were 28.2%, 92.5%, 33.6%, and 90.6% respectively. This is comparable with what was found in the index study. Patients may have difficulty interpreting the instructions verbally given on airway screening tests and the ULBT is not an exception to this. This difficulty can be easily overcome by demonstrating the tests to the patients. In studies comparing facial structure in African Americans versus Caucasians, African Americans have a broader nasal base, decreased nasal progression, bimaxillary protrusion, prominent lips and increased facial convexity.^{25,26} Asians have flatter malar prominence and mid face, more protuberant lips and more receded chin. This may make the occurrence of a positive upper lip bite test more common in their population. Hence, the differences in the findings in various populations may be

attributable to ethnic differences in craniofacial morphometry/ facial features as emphasised by anthropological literature.^{25,26}

Although the ratio of height to thyromental distance had the second highest sensitivity in this study, it only correctly identified about 35% of patients with DVL. This means that it mis-labelled nearly two-thirds of patients with DVL as being easy. Thus it is not very useful as a stand alone test in this population. Schmitt et al.⁷ originally described a RHTMD of ≥ 25 as being predictive of DVL and this was what was used in the index study. However, there seems to be a need to determine an optimal cut-off value for the RHTMD in various populations. Safavi et al.²⁷ determined, by discrimination analysis, the optimal cut-off value for their population at a RHTMD of 21.06 (with a sensitivity of 75.6% and specificity of 58.5%). Shah et al.⁸ used a cut-off point of 23.5 in their study. The lack of uniformity of cut-off values across populations may account for the differences in findings. In this population, with the ROC, at a cut-off value for the RHTMD of 24.7 (allowing for false positives of 5% or less), the sensitivity of the RHTMD was still low at 35% (Supplementary Table I).

The sensitivity of the thyromental distance was very low in this study (29.4%). This seems to imply that if the thyromental distance were to be used alone in this population, it would also miss the majority of patients with DVL resulting in an unacceptable number of unidentified DVL cases, compromising safety. This sensitivity was higher than what was reported by Merah et al.⁵ (15.4%). Shah et al.⁸ likewise found a very low sensitivity of the TMD (7.46%). Different cut-off values have been advocated for the TMD and variations in findings in different studies may be due, not only to anthropological differences but also a lack of standardisation of cut-off values and technique. Also, for the measurement of TMD, rigid rulers have been used in some studies.^{5,8,18} The use of a tape measure may, however, be more practicable for day to day use.

The MMT had the highest sensitivity but still missed more than 40% of DVL cases. There was no patient who had a Mallampati Grade of IV in this study. Differences were noted in patients' ability to follow the instructions for this test completely, but this was reduced as much as possible by demonstrating the test to the patients. While the specificity of the modified Mallampati was the lowest in this study, it was still high at 86.4%. Thus, it identified the majority of patients with EVL correctly. However, it had a very low PPV of 25%. Hence, we must ask, does a positive modified Mallampati test truly enhance our ability to make a correct diagnosis of DVL? Over-preparing for an easy intubation mislabelled as DVL, while potentially consuming resources, does not compromise patient safety. However, a problem might arise if anaesthetists, having lost confidence in its mis-labellings of EVL as DVL, ignore its warning signs, potentially missing even the correctly predicted patients with DVL. Mashour et al.²⁸ suggested that neck extension while performing the MMT may increase the specificity and predictive value of this tool. In their study, neck extension maintained the same sensitivity as without extension (83%) but the specificity was increased to 80% (with neck extension) from 70% (without neck extension). In the

index study, with the MMT performed with the head in neutral position, the specificity of the MMT of 86.4% was higher than what was found by Mashour et al.²⁸ even with neck extension. Perhaps, variations in findings of usefulness of the MMT are due to a lack of standardisation of technique as well as inter-observer variability.

The inter-incisor gap (IIG) had a low sensitivity of 11.8% even though it had a high specificity (97.0%). The use of calipers limited the probability that endpoints would vary with various patients. While Merah et al.⁵ found a higher sensitivity of 30.8%, this was still low. The sensitivity of the IIG in the present study was comparable with what was observed in a population of Indians (13.43%).⁸ Savva et al.²⁹ also found the IIG a non-useful tool in the prediction of DVL. Hence, it appears to be a common finding across populations that this test is not a very useful stand-alone test in prediction of DVL. It may be more useful as part of a multivariate assessment.

It is interesting to note that in spite of the seemingly low sensitivity of the ULBT and TMD in our study, these tests had the highest positive likelihood ratio of 11.71. This may indicate that when these tests are positive for DVL in this population based on the prevalence of 7.9%, there is a > 50% probability that the patient actually has DVL.²¹ The likelihood ratios for a positive test found in the study by Khan et al.⁶ and Eberhart et al.²⁴ were 3.78 and 6.76 respectively. The LR+ for the RHTMD was 4.68, which indicates only a small probability that a patient with an RHTMD predicting difficulty in this population would actually have DVL.²¹ However, at the cut-off value of 24.7 for a positive RHTMD using ROC in this study, the LR+ would be 7.02, indicating a moderate probability of DVL being present. None of the tests showed a high probability of being able to rule out DVL when absent based on the negative likelihood ratios.

To add some robustness to the findings, we used the data obtained to evaluate the degree of agreement between the tests. We found a high degree of agreement (73.4%) between the TMD and RHTMD ($K = 0.734$, $p < 0.001$) in this study. This is not surprising as the RHTMD is a test that incorporates the TMD. The lowest levels of agreement were between the ULBT and the RHTMD/TMD (see Supplementary Table II).

Age was an independent predictor of DVL in this study even though none of the study subjects was above 65 years. The incidence of DVL was higher in the middle-aged population than in the younger patients. The degenerative changes that occur in the cervical spine with increasing age, limiting neck mobility, may account for this.

Should we then say predictive airway assessment tests should be stopped based on the relative rarity of DVL and the poor and inconsistent results across various studies? Yentis³⁰ questioned the need and, indeed, usefulness of pre-anaesthetic airway assessment. He nonetheless concluded that it should be done, if only to force the anaesthetist into thinking about the airway. The authors of the index study agree that preoperative airway assessment should be done for all patients presenting for

anaesthesia to minimise the incidence of unanticipated difficult intubation. However, airway assessment must go beyond carrying out a series of tests as none of the tests studied was fool proof. Combinations of tests to further improve sensitivity should be considered, even though this may be at the expense of specificity. Beyond these, a high index of suspicion for DVL, detailed history and examination (including airway assessment tests) are crucial in planning the techniques of minimising difficulty and ensuring safe anaesthesia.

A limitation of this study was the small sample size considering the low prevalence of DVL which is responsible for the large confidence intervals seen. A study with a larger sample size in this population may further confirm the generalisability of these findings. The wide range of exclusion criteria may also have excluded some patients (for instance those older than 65 years) that could have added more information to the study. Another limitation was the fact that more than one person performed the Cormack and Lehane grading into EVL or DVL increasing the risk of inter-observer variability. The potential for this was reduced by limiting the anaesthetists performing the intubation to four experienced anaesthetists and also explaining the protocol, including visual aid reminders of the Cormack and Lehane grading to them prior to commencement of the research.

The variations in test performances noted in various populations, such as ours compared with Caucasian populations referenced, still support the hypothesis that performance of airway screening tests may be significantly affected by the anthropomorphic features of the population. It supports future work exploring the impact of anthropomorphic features on airway assessment. It increases understanding of a factor that appears key to variability of test performance between studies – population anthropometry.

Conflicts of interest

None declared

Acknowledgements

The authors wish to acknowledge the immeasurable contributions of Dr Ayodeji Isinkaye to the data analysis of this work.

References

1. Practice guidelines for the management of the difficult airway. An updated report by the American Society of Anesthesiologists Task Force on management of the difficult airway. *Anesthesiology* 2013;118:251-270.
2. Caplan RA, Posner KL, Ward RJ, Cheney FW. Adverse respiratory events in anaesthesia: A closed claims analysis. *Anesthesiology* 1990;72:828-833.
3. Benumof JL, Scheller MS. The importance of tracheal jet ventilation in the management of the difficult airway. *Anesthesiology* 1989;71:769-778.
4. Crosby ET, Cooper RM, Douglas MJ, Doyle DJ, Hung OR, Labrecque P, et al. The unanticipated difficult airway with recommendations for management. *Can J Anaesth* 1998;4:757-776.
5. Merah NA, Wong DT, Ffoulkes-Crabbe DJ, Kushimo OT, Bode CO. Modified Mallampati test, thyromental distance and inter-incisor gap are the best predictors of difficult laryngoscopy in West Africans. *Can J Anaesth* 2005;52:291-296.
6. Khan ZH, Kashfi A, Ebrahimkhani E. A comparison of the upper lip bite test (a simple new technique) with modified Mallampati classification in predicting difficulty in endotracheal intubation: a prospective blinded study. *Anesth Analg* 2003;96:595-599.

7. Schmitt HJ, Kirmse M, Radespiel-Troger M. Ratio of patient's height to thyromental distance improves prediction of difficult laryngoscopy. *Anaesth Int care* 2002;30:763-765.
8. Shah PJ, Dubey KP, Yadav JP. Predictive value of upper lip bite test and ratio of height to thyromental distance compared to other multivariate airway assessment tests for difficult laryngoscopy in apparently normal patients. *J Anaesth Clin Pharm* 2013;29:191-195.
9. Lundstrom LH, Vester Andersen M, Moller AM, Charuluxananan S, L'Hermite J, Wetterslev J. Poor prognostic value of the modified Mallampati score: a meta-analysis involving 177,088 patients. *Br J Anaesth* 2011;107:659-667.
10. Wong JK, Tongier WK, Armbruster SC, White PF. Use of the intubating laryngeal mask airway to facilitate awake orotracheal intubation in patients with cervical spine disorders. *J Clin Anesth* 1999;11:346-348.
11. Aziz MF, Dillman D, Fu R, Brambrink AM. Comparative effectiveness of the C-MAC video laryngoscope versus direct laryngoscopy in the setting of the predicted difficult airway. *Anesthesiology* 2012;116:629-636.
12. Enomoto Y, Asai T, Arai T, Kamishima K, Okuda Y. Pentax AWS, a new videolaryngoscope, is more effective than the Macintosh laryngoscope for tracheal intubation in patients with restricted neck movements: A randomized comparative study. *Br J Anaesth* 2008;100:544-548.
13. Jungbauer A, Schumann M, Brunkhorst V, Börgers A, Groeben H. Expected difficult tracheal intubation: A prospective comparison of direct laryngoscopy and video laryngoscopy in 200 patients. *Br J Anaesth* 2009;102:546-550.
14. Oates JDL, Macleod AD, Oates PD, Pearsall FJ, Howie JC, Murray GD. Comparison of two methods for predicting difficult intubation. *Br J Anaesth* 1991;66:305-309.
15. Randell I. Prediction of difficult intubation. *Acta Anaesthesiol Scand* 1996;40:1016-1023.
16. Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraska B, Fereiberger D. A clinical sign to predict difficult tracheal intubation: A prospective study. *Can Anaesth Soc J* 1985;32:429-434.
17. Patil VU, Stehling LC, Zauder HL. *Fiberoptic endoscopy in anaesthesia: visualizing the difficult airway*. Chicago: Yearbook Medical Publishers; 1983. p. 79-80.
18. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. *Anesthesiology* 2005;103:429-437.
19. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. *Br J Anaesth* 1988;61:211-216.
20. Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia* 1984;39:1105-1111.
21. Fagan TJ. Nomogram for Baye's theorem. *N. Engl J Med* 1975;293:257.
22. Rose DK, Cohen MM. The airway: Problems and predictions in 18,500 patients. *Can J Anaesth* 1994;41:372-383.
23. Jin Huh, Hwa-Yong Shin, Seong-Hyop Kim, Tae-Kyoon Yoon, Duk-Kyung Kim. Diagnostic predictor of difficult laryngoscopy: the hyomental distance ratio. *Anesth Analg* 2009; -108: -544-548.
24. Eberhart LH, Arndt C, Cierpka T, Schwanekamp J, Wulf H, Putzke C. The reliability and validity of the upper lip bite test compared with the Mallampati classification to predict difficult laryngoscopy: An external prospective evaluation. *Anesth Analg* 2005;101: 284-289.
25. Fang F, Clapham PJ, Chung KC. A systematic review of interethnic variability in facial dimensions. *Plast Reconstr Surg* 2011;127:874-881.
26. Farkas LG, Forest CR, Litsas L. Revision of neoclassical facial canons in young adult Afro Americans. *Aesthetic Plast Surg* 2000;24:179-184.
27. Safavi MR, Honarmand A, Zare N. A comparison of the ratio of height to thyromental distance with the modified Mallampati and the upper lip bite test in predicting difficult laryngoscopy. *Saudi J Anaesth* 2011;5:258-263.
28. Mashour GA, Sandberg WS. Craniocervical extension improves the specificity and predictive value of the Mallampati airway evaluation. *Anesth Analg* 2006;103:1256-1259.
29. Savva D. Prediction of difficult tracheal intubation. *Br J Anaesth* 1994;73:149-153.
30. Yentis SM. Predicting difficult intubation: Worthwhile exercise or pointless ritual? *Anaesthesia* 2002;57:105-109.