

# Airway Ultrasound as Predictor of Difficult Direct Laryngoscopy: A Systematic Review and Meta-analysis

Andrea Carsetti, MD,\*† Massimiliano Sorbello, MD,‡ Erica Adrario, MD,\*† Abele Donati, MD, PhD,\*† and Stefano Falcetta, MD†

**BACKGROUND:** Despite several clinical index tests that are currently applied for airway assessment, unpredicted difficult laryngoscopy may still represent a serious problem in anesthesia practice. The aim of this systematic review and meta-analysis was to evaluate whether preoperative airway ultrasound can predict difficult direct laryngoscopy in adult patients undergoing elective surgery under general anesthesia.

**METHODS:** We searched the Medline, Scopus, and Web of Science databases from their inception to December 2020. The population of interest included adults who required tracheal intubation for elective surgery under general anesthesia without clear anatomical abnormalities suggesting difficult laryngoscopy. A bivariate model has been used to assess the accuracy of each ultrasound index test to predict difficult direct laryngoscopy.

**RESULTS:** Fifteen studies have been considered for quantitative analysis of summary receiver operating characteristic (SROC). The sensitivity for distance from skin to epiglottis (DSE), distance from skin to hyoid bone (DSHB), and distance from skin to vocal cords (DSVC) was 0.82 (0.74–0.87), 0.71 (0.58–0.82), and 0.75 (0.62–0.84), respectively. The specificity for DSE, DSHB, and DSVC was 0.79 (0.70–0.87), 0.71 (0.57–0.82), and 0.72 (0.45–0.89), respectively. The area under the curve (AUC) for DSE, DSHB, DSVC, and ratio between the depth of the pre-epiglottic space and the distance from the epiglottis to the vocal cords (Pre-E/E-VC) was 0.87 (0.84–0.90), 0.77 (0.73–0.81), 0.78 (0.74–0.81), and 0.71 (0.67–0.75), respectively. Patients with difficult direct laryngoscopy have higher DSE, DSVC, and DSHB values than patients with easy laryngoscopy, with a mean difference of 0.38 cm (95% confidence interval [CI], 0.17–0.58 cm;  $P = .0004$ ), 0.18 cm (95% CI, 0.01–0.35 cm;  $P = .04$ ), and 0.23 cm (95% CI, 0.08–0.39 cm;  $P = .004$ ), respectively.

**CONCLUSIONS:** Our study demonstrates that airway ultrasound index tests are significantly different between patients with easy versus difficult direct laryngoscopy, and the DSE is the most studied index test in literature to predict difficult direct laryngoscopy. However, it is not currently possible to reach a definitive conclusion. Further studies are needed with better standardization of ultrasound assessment to limit all possible sources of heterogeneity. (Anesth Analg 2021;00:00–00)

## KEY POINTS

- **Question:** Is preoperative upper airway ultrasound able to predict a difficult airway in adult patients undergoing elective surgery under general anesthesia without clear anatomical abnormalities suggesting difficult laryngoscopy?
- **Findings:** The distance from skin to epiglottis was the most extensively assessed index test in literature and seems accurate to predict difficult laryngoscopy.
- **Meaning:** The high heterogeneity performing ultrasound and the limited number of studies do not allow to reach a definitive conclusion, and the routine use of ultrasound to predict difficult laryngoscopy cannot still be recommended.

## GLOSSARY

? = uncertain risk of bias; + = low risk of bias; – = high risk of bias; **AUC** = area under the curve; **BURP** = backward-upward-rightward-pressure; **CI** = confidence interval; **CL** = Cormack-Lehane; **DSE** = distance from the skin to epiglottis; **DSHB** = distance from skin to the hyoid bone; **DSVC** = distance from skin to vocal cords; **ENT** = eye-nose-throat; **FN** = false negative; **FP** = false positive; **GRADE** = Grading of Recommendations, Assessment, Development and Evaluation; **HMD-E** = hyomental distance in extended neck position; **HMD-N** = hyomental distance in neutral neck position; **HSROC** = hierarchical summary receiver operating characteristic; **IV** = inverse variance; **LR+** = positive likelihood ratio; **LR–** = negative likelihood ratio; **MDs** = mean differences; **NPV** = negative predictive value; **POGO** = percentage of glottic opening; **PPV** = positive predictive value; **Pre-E/E-VC** = ratio between the depth of the pre-epiglottic space and the distance from the epiglottis to the vocal cords; **PRISMA-DTA** = Preferred Reporting Items for Systematic Reviews and Meta-analyses of diagnostic test accuracy studies; **PRISMA-P** = Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols; **PROSPERO** = International Prospective Register of Systematic Reviews; **QUADAS-2** = quality assessment of diagnostic accuracy studies; **ROC** = receiver operating characteristic; **SD** = standard deviation; **SENS** = sensitivity; **SPEC** = specificity; **SROC** = summary receiver operating characteristic; **TN** = true negative; **TP** = true positive; **UA-US** = upper airway ultrasound; – = high risk of bias

Unsuccessful airway management leads to serious morbidity and mortality, and the unanticipated difficult intubation is a potentially life-threatening event during anesthesia. Several bedside screening tests are used in clinical practice to identify patients at the risk of difficult airway. Despite their accuracy and benefit were well proven in the literature and daily practice, a small number of patients classified to an easy airway may still present an unexpected difficulty. Predicting a “difficult airway” is not at all an easy task for all patients<sup>1-7</sup>: many structures and functional units are involved in the pathogenesis of a difficult airway, which is a dynamic phenomenon and highly dependent on the operator’s experience. We then need to consider that many studies have been performed with different definitions and criteria for difficult airway, including the interobserver variability during the assessment, given that not all measurements are provided with objective parameters. Finally, we need to consider that all the factors involved in the genesis of a difficult airway may be differently combined in a large number of possibilities.<sup>8</sup> This means that predicting a difficult airway represents the attempt to adopt a quantitative assessment of many qualitative and quantitative parameters, with the final conclusion that “any difficult airway is difficult its own way.”<sup>9</sup> This may also explain why the incidence of difficult airway and difficult intubation varies from 5% to 22%,<sup>10-12</sup> with important implications for clinical practice and patients’ outcomes. Nowadays, several clinical tests recommended by current guidelines for airway assessment<sup>13</sup> make patients with difficult airways easily identifiable. On the other hand, a minority of subjects classified with easy airways will be instead unexpectedly difficult to manage.<sup>9</sup> Hence, the need to develop adequate

tools to successfully predict not a difficult airway but an unexpectedly difficult airway in patients previously classified as easy, possibly including in the clinical evaluation some objective index tests to increase sensibility and specificity and reduce interobserver variability. Although this may seem a problem interesting to a very small number of patients, it may have serious life-threatening consequences when it occurs.

For many years, ultrasounds have been used as a complementary tool to predict difficulty in airway management, both from a qualitative and quantitative perspective.

To date, many studies have been published with the aim to find an effective ultrasound indicator to predict a difficult airway, but with significant limitations due to large variability of the sample homogeneity, to the kind of population included, and to the absence of a standardized protocol for ultrasound assessments.

The aim of this systematic review and meta-analysis was to evaluate whether preoperative upper airway ultrasound (UA-US) can predict a difficult airway in adult patients undergoing elective surgery under general anesthesia without clear anatomical evidence of difficult airway on standard clinical examination. Moreover, the mean difference (MD) of UA-US index tests between patients with easy and difficult direct laryngoscopy has been investigated.

## METHODS

### Protocol and Guidance for Conducting and Reporting

The study protocol (Supplemental Digital Content 1, Supplemental Material, <http://links.lww.com/AA/D759>) was conducted following Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) guidelines.<sup>14</sup> The methodology for conducting and reporting the systematic review followed Preferred Reporting Items for Systematic Reviews and Meta-analyses of diagnostic test accuracy studies (PRISMA-DTA) guidelines.<sup>15</sup> The protocol has been registered on International Prospective Register of Systematic Reviews (PROSPERO) (CRD42020156134).

### Eligibility Criteria

We considered adult patients (age  $\geq 18$  years old) who required tracheal intubation for elective surgery under general anesthesia. Criteria for inclusion of a retrieved paper were that it had studied only patients without clear anatomical abnormalities suggesting difficult intubation. Patients with a history of previous difficult intubation or expected difficult laryngoscopy have been excluded. The intervention included preoperative UA-US assessment of neck soft tissue. No US index tests’ selection has been made a priori. Each study assessing difficult laryngoscopy with the US has been

From the \*Department of Biomedical Sciences and Public Health, Università Politecnica delle Marche, Ancona, Italy; †Anesthesia and Intensive Care Unit, Azienda Ospedaliero Universitaria Ospedali Riuniti, Ancona, Italy; and ‡Anesthesia and Intensive Care, Policlinico San Marco University Hospital, Catania, Italy.

Accepted for publication September 22, 2021.

Funding: None.

Conflicts of Interest: See Disclosures at the end of the article.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal’s website ([www.anesthesia-analgesia.org](http://www.anesthesia-analgesia.org)).

Study registration: PROSPERO: CRD42020156134 [https://www.crd.york.ac.uk/prospero/display\\_record.php?ID=CRD42020156134](https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020156134).

Reprints will not be available from the authors.

Address correspondence to Andrea Carsetti, MD, Department of Biomedical Sciences and Public Health, Università Politecnica delle Marche, via Tronto 10/A, 60126 Ancona, Italy. Address e-mail to [a.carsetti@univpm.it](mailto:a.carsetti@univpm.it).

Copyright © 2021 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Anesthesia Research Society. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1213/ANE.0000000000005839

considered, despite the index test applied. Examples of UA-US index tests include (but are not limited to): distance from skin to epiglottis (DSE), distance from skin to the hyoid bone (DSHB), distance from skin to vocal cords (DSVC), and the ratio between the depth of the pre-epiglottic space and the distance from the epiglottis to the vocal cords (Pre-E/E-VC). The primary outcome was the predictive value of neck US index tests to anticipate difficult direct laryngoscopy. The secondary outcome was to determine the MD of UA-US index tests between patients with easy and difficult direct laryngoscopy. Direct laryngoscopy view was classified according to Cormack-Lehane (CL) grade.<sup>16</sup> CL I-II were considered as easy laryngoscopy, whereas CL III-IV were considered as difficult laryngoscopy. According to modified CL classification,<sup>17</sup> IIb grade was also considered as difficult laryngoscopy.

Eligible studies were observational trials. We excluded conference proceedings, abstracts, and studies conducted on animals. Even if some bias coming from language restriction cannot be excluded for diagnostic test assessment systematic reviews, only studies published in English were considered.

### Search Strategy

We searched the Medline, Scopus, and Web of Science databases from their inception to December 2020. We combined the terms “airway ultrasound,” “neck ultrasound,” “difficult laryngoscopy,” “difficult airways,” and “difficult intubation” (Supplemental Digital Content 1, <http://links.lww.com/AA/D759>). Reference lists of eligible studies and review articles have been assessed.

### Study Selection and Data Extraction

Two researchers (A.C. and S.F.) independently screened the titles and abstracts of all papers resulting from the database search. Subsequently, they independently assessed the full text of the papers selected from titles’ and abstracts’ screening. The same investigators independently performed data extraction. Any discrepancies that arose during the selecting process and data extraction were solved by consensus or by the decision of a third independent researcher (A.D.).

### Assessment of Risk of Bias and Quality of the Evidence

Two trained investigators (A.C. and S.F.) independently rated the quality of the selected studies. As per the Cochrane DTA handbook,<sup>18</sup> the quality assessment of diagnostic accuracy studies (QUADAS-2) tool was used to assess for risk of bias and applicability concerns in patients’ selection, index test, reference standard, and flow and timing. Each item was evaluated as low, unclear, or high risk of bias.<sup>19</sup> The highest risk of bias shown for any item was used to

determine the overall risk of bias for the study. The overall quality of the evidence for the primary outcome has been assessed according to the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.<sup>20,21</sup>

### Statistical Analysis

The statistical analyses were performed using metandi and midas in STATA (StataCorp 2021; Stata Statistical Software: Release 17; StataCorp LLC) and RevMan, version 5.3 (Cochrane Collaboration). The bivariate model proposed by Reitsma et al<sup>22</sup> has been used to assess the accuracy of each US index test to predict difficult direct laryngoscopy. In the absence of covariates, this model also gives hierarchical summary receiver operating characteristic (HSROC) parameters as it is equivalent to the HSROC analysis proposed by Rutter and Gatsonis<sup>23</sup> and Harbord et al.<sup>24</sup> Only US index tests assessed by  $\geq 5$  studies were considered for quantitative summary receiver operating characteristic (SROC) analysis.<sup>25</sup> In the presence of an appropriate number of studies, a subgroup analysis was considered to investigate the potential sources of heterogeneity.

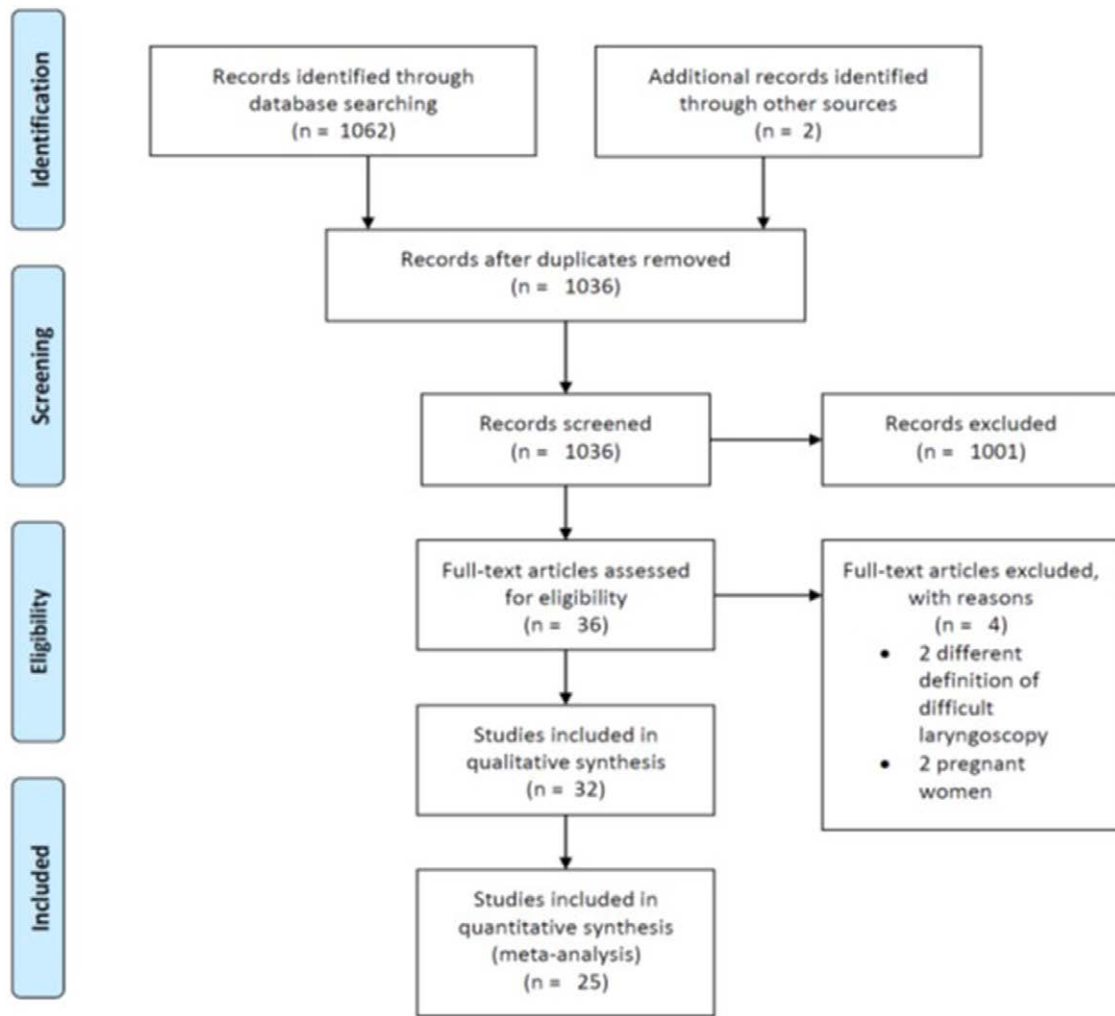
The difference in US index tests between the 2 groups was analyzed using the inverse variance random-effects model and was expressed as MDs. A *P* value of  $< .05$  was considered statistically significant. Heterogeneity was assessed using the  $X^2$  test and the  $I^2$  test, with  $I^2 > 50\%$  being considered substantial.<sup>26</sup> The possibility of publication bias was assessed by the visual estimate of funnel plot and by the regression test of Egger test and Begg test when  $\geq 10$  trials were pooled.<sup>27</sup>

## RESULTS

### Study Selection and Study Characteristics

One thousand sixty-four titles were retrieved: after removal of duplicates, we screened the titles/abstracts of 1036 records and assessed the full text of 36 articles. Finally, 32 studies were included (Figure 1) enrolling a total of 6881 patients. The studies were published from 2003 to 2020. The main characteristics of the selected studies are described in Supplemental Digital Content 1, Tables 1 and 2, <http://links.lww.com/AA/D759>.

All studies considered adult patients undergoing general anesthesia for elective surgery, excluding patients with predicted difficult airway management basing on medical history or anatomical abnormalities. Six studies selectively enrolled obese patients,<sup>28–33</sup> while 5 studies excluded this population.<sup>34–38</sup> Almost all studies defined the CL grade III and IV as difficult direct laryngoscopy. Only 4 studies included also CL grade IIb.<sup>39–42</sup> Some variability was observed regarding head positioning for UA-US evaluation, ranging from the neutral, extended, or sniffing position. A further source of potential heterogeneity was the application of external laryngeal manipulation during



**Figure 1.** PRISMA flow diagram. PRISMA indicates Preferred Reporting Items for Systematic Review and Meta-Analysis.

laryngoscopy. The backward-upward-rightward-pressure (BURP) maneuver was applied in 3 studies,<sup>28,40,43</sup> 12 did not use it,<sup>29,32,34,37,44–51</sup> and 17 did not mention if any external laryngeal manipulation was applied to improve laryngeal view.<sup>30,31,33,35,36,38,39,41,42,52–59</sup> Some studies reported sensitivity, specificity, and area under the curve receiver operating characteristic (AUC-ROC) for each UA-US index test, whereas other studies only compared the mean values of UA-US index test between the patients' groups (Supplemental Digital Content 1, Table 1, <http://links.lww.com/AA/D759>).

A variety of different UA-US measurements were evaluated by different studies for their effectiveness in predicting difficult direct laryngoscopy (Supplemental Digital Content 1, Table 1, <http://links.lww.com/AA/D759>).

#### Excluded Studies

Two studies were excluded because of a different definition of difficult direct laryngoscopy,<sup>60,61</sup> and 2 studies enrolling pregnant women were excluded.<sup>62,63</sup>

#### Risk of Bias and Quality of Evidence

The quality assessment of the studies is summarized in Table 1. Almost all studies clearly stated that US assessment was performed before surgery, and a different physician performed laryngoscopy blindly to US result. The expertise of physicians who performed US and laryngoscopy was not always clearly stated. Two studies did not clearly state exclusion criteria.<sup>30,42</sup>

The overall quality of evidence was low/very low due to the high heterogeneity between studies about the classification of difficult direct laryngoscopy, BURP application, and patient head position during US measurement (Supplemental Digital Content 1, Tables 4–7, <http://links.lww.com/AA/D759>).

#### Primary Outcome

Fifteen studies have been considered for quantitative analysis of SROC. DSE, DSHB, DSVC, and Pre-E/VC were the most extensively reported index tests (Supplemental Digital Content 1, Table 2, <http://links.lww.com/AA/D759>). Individual study

**Table 1. Assessment of Risk of Bias According to QUADAS-2**

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Ezri et al (2003) <sup>28</sup>	+	+	?	?	-	+	+
Komatsu et al (2007) <sup>29</sup>	+	+	?	?	-	+	+
Adhikari et al (2011) <sup>59</sup>	+	?	?	?	+	+	+
Wojtczak (2012) <sup>30</sup>	?	?	?	?	-	+	+
Hui and Tsui (2014) <sup>52</sup>	+	+	+	+	+	+	+
Wu et al (2014) <sup>46</sup>	+	+	+	?	+	+	+
Andruszkiewicz et al (2016) <sup>53</sup>	+	+	+	+	+	+	+
Pinto et al (2016) <sup>34</sup>	+	?	?	?	+	+	+
Reddy et al (2016) <sup>54</sup>	+	?	+	+	+	+	+
Soltani Mohammadi et al (2016) <sup>36</sup>	+	+	?	+	+	+	+
Parameswari et al (2017) <sup>35</sup>	+	+	+	+	+	+	+
Yao et al (2017) <sup>55</sup>	+	+	+	+	+	+	+
Yao and Wang (2017) <sup>43</sup>	+	+	+	+	+	+	+
Chan et al (2018) <sup>39</sup>	+	+	+	+	+	+	+
Falcetta et al (2018) <sup>40</sup>	+	+	+	+	+	+	+
Petrisor et al (2018) <sup>31</sup>	+	?	+	+	+	+	+
Rana et al (2018) <sup>47</sup>	+	+	+	+	+	+	+
Yilmaz et al (2018) <sup>32</sup>	+	+	+	+	-	+	+
Abo Sabaa et al (2019) <sup>48</sup>	+	?	?	?	+	+	+
Alessandri et al (2019) <sup>41</sup>	+	?	+	+	+	+	+
Fulkerson et al (2019) <sup>49</sup>	+	+	+	+	+	+	+
Koundal et al (2019) <sup>37</sup>	+	+	+	+	+	+	+
Wang et al (2019) <sup>56</sup>	+	+	+	+	+	+	+
Yadav et al (2019) <sup>50</sup>	+	?	?	+	+	+	+
Abdelhady et al (2020) <sup>38</sup>	+	+	+	+	+	+	+
Daggupati et al (2020) <sup>51</sup>	+	+	+	+	+	+	+
Martínez-García et al (2021) <sup>44</sup>	+	+	+	?	+	+	+
Ni et al (2020) <sup>45</sup>	+	+	+	+	+	+	+
Petrisor et al (2020) <sup>57</sup>	+	+	+	+	+	+	+
Sharma and Bhalla (2020) <sup>33</sup>	+	+	+	+	-	+	+
Shetty and Smruthi (2020) <sup>58</sup>	+	+	+	+	+	+	+
Senapathi et al (2020) <sup>42</sup>	?	?	?	+	?	+	+

Abbreviations: ?, uncertain risk of bias; +, low risk of bias; -, high risk of bias; QUADAS-2, quality assessment of diagnostic accuracy studies.

sensitivity, specificity, AUC, and cutoff of each US index test are presented in Supplemental Digital Content 1, Table 3, <http://links.lww.com/AA/D759>. The overall sensitivity, specificity, and positive and negative likelihood ratios for DSE, DSHB, DSVC, and Pre-E/E-VC are presented in Table 2. The forest plot for DSE showing the individual study sensitivity and specificity and the SROC curve are shown in Figure 2, while the forest plot and SROC curves for DSHB, DSVC, and Pre-E/E-VC are shown in the supplemental material (Supplemental Digital Content 1, Figures 1–3, <http://links.lww.com/AA/D759>). The AUC for DSE, DSHB, DSVC, and Pre-E/E-VC was 0.87 (0.84–0.90), 0.77 (0.73–0.81), 0.78 (0.74–0.81), and 0.71 (0.67–0.75), respectively (Table 2). HSROC parameters for each US measurement were reported in Supplemental Digital Content 1, Table 8, <http://links.lww.com/AA/D759>.

SROC analysis for the other UA-US index tests was not run because data could not be obtained for ≥5 studies. Subgroup and sensitivity analyses to identify the source of heterogeneity were not possible due to the limited number of study for each index test.

**Secondary Outcome**

Seventeen studies have been considered for secondary outcome.<sup>28–33,38,41,44,46,48–51,53,59,64</sup> Patients with difficult direct laryngoscopy had higher DSE, DSVC, and DSHB values than patients with easy laryngoscopy, with MD of 0.38 cm (95% confidence interval [CI], 0.17–0.58 cm;  $P = .0004$ ), 0.18 cm (95% CI, 0.01–0.35 cm;  $P = .04$ ), and 0.23 cm (95% CI, 0.08–0.39 cm;  $P = .004$ ), respectively (Figure 3; Supplemental Digital Content 1, Figures 4 and 5, <http://links.lww.com/AA/D759>). However, the level of heterogeneity was very high. Hyomental distance in neutral neck position (HMD-N) and hyomental distance in extended neck position (HMD-E) were significantly shorter in patients with difficult laryngoscopy (MD for HMD-N, -0.33 cm [95% CI, -0.43 to -0.19 cm],  $P < .00001$ ,  $I^2 = 0%$ ,  $P = .55$ ; MD for HMD-E, -0.60 cm [95% CI, -0 to 92; -0.28 cm],  $P = .0003$ ,  $I^2 = 83%$ ,  $P = .0001$ ) (Supplemental Digital Content 1, Figures 6 and 7, <http://links.lww.com/AA/D759>).

**DISCUSSION**

Prediction of difficult airways is a crucial aspect of anesthesia management. Several studies have been

**Table 2. Diagnostic Test Accuracy Results**

Index test	Sensitivity	Specificity	LR+	LR-	AUC
DSE	0.82 (0.74–0.87)	0.79 (0.70–0.87)	3.91 (2.65–5.76)	0.23 (0.16–0.33)	0.87 (0.84–0.90)
DSHB	0.71 (0.58–0.82)	0.71 (0.57–0.82)	2.46 (1.50–4.04)	0.40 (0.25–0.66)	0.77 (0.73–0.81)
DSVC	0.75 (0.62–0.84)	0.72 (0.45–0.89)	2.63 (1.16–5.98)	0.36 (0.20–0.62)	0.78 (0.74–0.81)
Pre-E/E-VC	0.65 (0.22–0.93)	0.68 (0.43–0.85)	2.02 (1.0–4.07)	0.51 (0.63–6.11)	0.71 (0.67–0.75)

Data reported as estimate value (95% CI).

Abbreviations: AUC, area under the curve; CI, confidence interval; DSE, distance from skin to epiglottis; DSHB, distance from skin to hyoid bone; DSVC, distance from skin to vocal cords; E-VC, distance from the epiglottis to the midpoint of the distance between the vocal cords; LR+, positive likelihood ratio; LR-, negative likelihood ratio; Pre-E, pre-epiglottis space.

conducted aiming to determine the role of UA-US to predict easy versus difficult direct laryngoscopy in patients without clear evidence of difficulties at standard clinical evaluation. To our knowledge, this is the first diagnostic test accuracy systematic review and meta-analysis aiming to understand the US ability to predict difficult laryngoscopy in this clinical scenario, clarifying the possible role of this tool in clinical practice. Among the several UA-US index tests considered in the literature, we showed that DSE seems accurate to predict difficult laryngoscopy in this population with AUC-SROC of 0.87 (95% CI, 0.84–0.90) (Table 2). This result comes from 10 studies considering 1812 patients.<sup>35,37,38,40,42,44,46,50,51,58</sup> Particularly, DSE was higher in patients who showed higher CL grades. According to the prevalence of difficult direct laryngoscopy reported by considered studies, the positive predictive value (PPV) ranged from 30.26% to 49.4%, while the negative predictive value (NPV) ranged from 94.61% to 97.53% (Supplemental Digital Content 1, Table 4, <http://links.lww.com/AA/D759>). This means that despite the good sensitivity and specificity, DSE is probably most useful in case of a negative result to identify a patient who will not present a difficult airway (when the test is negative, the probability of an easy laryngoscopy is about 95%–97%). This is due to the low prevalence of patients with a difficult airway in the target population (ranging from 10% to 20% in the included studies). On the other hand, a positive result suggests a probability of about 30% to 50% that the patient will be difficult to intubate. In clinical practice, as missing a difficult airway may be dangerous, anesthesiologist should approach patients with a positive test with caution (eg, using videolaryngoscopy), as some of these will be really difficult. “Overtreating” in this setting may be prudential.

DSE was assessed using a linear probe placed in the transverse plane and measuring the thickness of the pre-epiglottic space at the midline (Supplemental Digital Content 1, Figure 8, <http://links.lww.com/AA/D759>). Two studies were not included in quantitative analysis as they used different methods for this evaluation (average of midline and lateral measurements<sup>34</sup> and lateral parasagittal scanning<sup>45</sup>). Despite this technical difference, they too found that a higher value of DSE was significantly associated

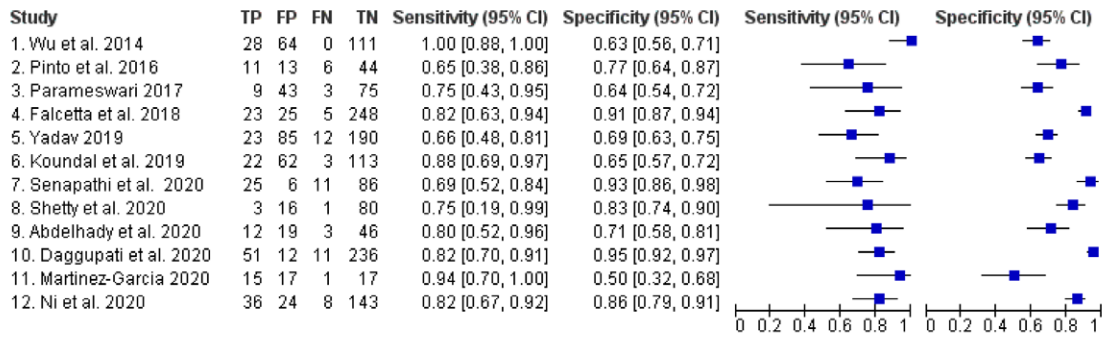
with difficult laryngoscopy. DSHB, DSVC, and Pre-E/E-VC were other UA-US index tests frequently assessed. Furthermore, the 95% prediction regions for each index test were quite wide and did not permit a definitive conclusion, including the identification of which index test may be significantly better than the others. This derives from the small number of studies that could be included in the quantitative analysis.

Several different index tests have been assessed by single studies, not allowing for an overall quantitative analysis (Supplemental Digital Content 1, Table 1, <http://links.lww.com/AA/D759>). The hyomen-tal distance measured in the extended head position was significantly shorter in patients with difficult direct laryngoscopy,<sup>30,53</sup> and it showed to be potentially useful to predict difficult laryngoscopy both in obese patients and general population, with the AUC of 0.87 (cutoff <5.5 cm)<sup>31</sup> and 0.85, respectively (cutoff <5.6 cm).<sup>48</sup> Conversely, tongue thickness showed a lower capacity to discriminate between the 2 groups of patients, with AUC ranging from 0.56 (95% CI, 0.24–0.88) to 0.72 (95% CI, 0.62–0.81).<sup>43,50,58</sup>

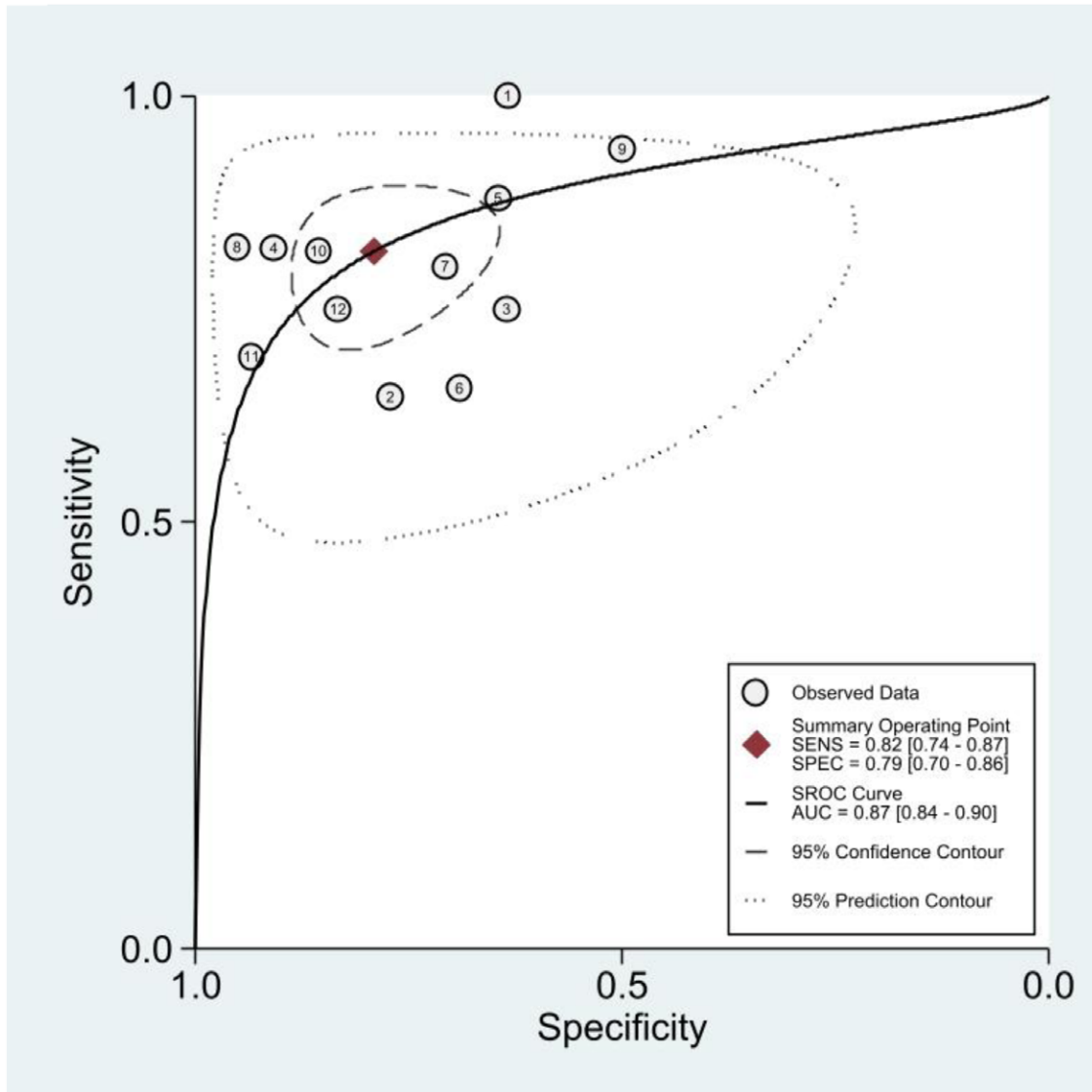
The cutoff value of UA-US index tests to discriminate between patients with possible easy and difficult laryngoscopy was different between studies for the same index test. For DSE, the cutoff values ranged from 1.615 to 2.75 cm. SROC can only show how sensitivity and specificity vary changing the threshold. Even if the optimal cutoff has not been statistically defined by our analysis, the best predicting performance has been shown by 3 large studies<sup>40,45,51</sup> enrolling 822 patients considering a similar threshold (>2–2.5 cm). This cutoff value seems reasonable for clinical application. On the contrary, the worse predicting value for DSE was reported by Alessandri et al<sup>41</sup> (AUC, 0.644 [95% CI, 0.54–0.749]) in 194 patients undergoing eye-nose-throat (ENT) surgery. However, the study, unlike others, included a specific patient’s population (ENT surgery) and did not report the cutoff value, and it was not possible to extrapolate full data to include the study in SROC analysis.

Although our results showed that US index tests (particularly DSE) are probably useful to predict difficult direct laryngoscopy, they need to be interpreted with caution due to several limitations of the available evidence.<sup>65</sup> A high clinical and methodological

**A**



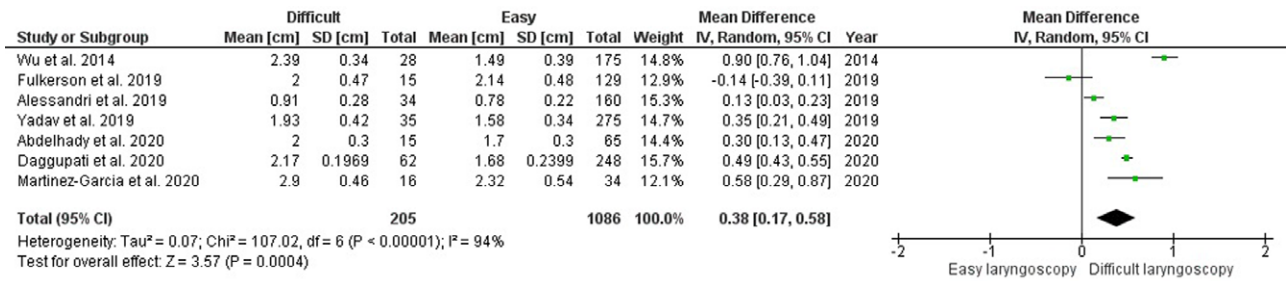
**B**



**Figure 2.** Accuracy for DSE. A, Forest plot for DSE showing the individual study sensitivity and specificity. B, SROC curve for DSE. AUC indicates area under the curve; CI, confidence interval; DSE, distance from skin to epiglottis; FN, false negative; FP, false positive; SENS, sensitivity; SPEC, specificity; SROC, summary receiver operating characteristic; TN, true negative; TP, true positive.

heterogeneity has been found between studies. It may be probably due to the patients' selection and some differences to perform US assessment (eg, head position). External laryngeal manipulation (BURP

maneuver) was not uniformly applied, and several studies did not mention if it was allowed during direct laryngoscopy. This aspect may affect the results as BURP may significantly improve laryngeal



**Figure 3.** Forest plot for the mean difference of DSE between patients with easy and difficult direct laryngoscopies. CI indicates confidence interval; DSE, distance from skin to epiglottis; IV, inverse variance; SD, standard deviation.

visualization and consequently CL classification. Unfortunately, the limited number of studies did not allow further analysis to investigate the source of heterogeneity.

The current guidelines for preprocedural evaluation recommend using a combination of the validated tests to predict difficult airway management as no single predictive sign is sufficient by itself.<sup>13</sup> In this contest, the integration of US airway assessment with routinely used tests should be investigated to clarify the potential role of this technique in the frame of preprocedural patients' evaluation.<sup>8</sup> Routine airway clinical assessment is limited to external suprahyoid evaluation (ie, mouth opening, tongue wideness, interincisor distance, and hyomental and thyromental distances). However, US is able to assess subhyoid parameters that could help us to identify patients with normal routine clinical assessment but potentially difficult airways due to the thickness of soft tissues. Considering a diagnostic test, a proper patients' selection is needed to optimize the posttest probability. In fact, as previously discussed, applying the UA-US as a routine assessment will reduce the PPV of the test as the overall prevalence of difficult laryngoscopy in the general population is low. Thus, we believe that DSE >2 to 2.5 cm may have a role in case of doubt for potential difficulties after considering the other tests routinely applied. Although difficult airway may be easily identified when multiple standard tests are positive or in the presence of evident anatomical abnormalities, a "gray zone" of uncertainty may still be present when only a few clinical signs of difficulty can be shown. In this situation, DSE may help to reasonably rule out the difficult airway when negative or suggest a prudent approach in case of positivity.

We believe that UA-US may represent a powerful tool to improve the performance of difficult airway management predictive tests, providing an objective assessment of specific index tests, thus restricting the interobserver variability. On the other hand, to reach this goal, we need well-designed and conducted studies with standardization of UA-US protocols, with univocal and common definitions of difficult airways (as, eg, adoption of the modified CL classification<sup>17</sup> or

the percentage of glottic opening (POGO) score<sup>66</sup> or Freemantle score<sup>67</sup> if videolaryngoscopes are used) to discriminate between easy and difficult direct laryngoscopies, inclusion of adjuncts such as tracheal introducer<sup>68</sup> or external maneuver and extension of airway evaluation with UA-US to assess also rescue options such as supraglottic airway placement<sup>69,70</sup> or emergency front of neck performance.

Our study has several limitations. First, the evidence supporting the utility of UA-US to predict difficult laryngoscopy is weak due to the limited number of studies included in the quantitative analysis, giving wide CIs that do not permit a definitive conclusion. Second, we were not able to perform an advanced analysis about the source of heterogeneity due to the lack of a sufficient number of studies. Several sources of heterogeneity may be supposed, but not currently supported by rigorous analysis. However, we think that our study highlights the need to expand and standardize the research in this field. Third, an optimal cutoff has not been statistically defined, and further research is needed to definitively respond to this question. Finally, we decided to include studies that assessed the usefulness of UA-US in the setting of elective surgery. However, the UA-US may have an important role also in the management of emergencies. Airway assessment is fundamental in any situation, and sometimes some routine tests are not applicable due to a lack of patient collaboration. In this contest, UA-US may be extremely helpful as an objective assessment. This aspect needs to be further investigated.

In conclusion, our study demonstrates that airway UA-US index tests are significantly different between patients with easy versus difficult direct laryngoscopy, and the DSE is the most studied index test in literature to predict difficult direct laryngoscopy. It may help to rule out the probability of a true difficult laryngoscopy in a selected population with uncertain difficult airways based on the clinical assessment. However, it is not currently possible to reach a definitive conclusion. Further studies are needed with better standardization of US assessment to limit all possible sources of heterogeneity before recommending routine use of



UA-US as a decisional support tool during periprocedural airway assessment, and the improvement of difficult airway prediction based on both UA-US and standard routine assessment needs to be further investigated. ■■

## DISCLOSURES

**Name:** Andrea Carsetti, MD.

**Contribution:** This author helped perform literature search, review and perform data analysis, and draft the manuscript.

**Conflicts of Interest:** None.

**Name:** Massimiliano Sorbello, MD.

**Contribution:** This author helped with substantial manuscript revision.

**Conflicts of Interest:** M. Sorbello has received paid consultancy from Teleflex Medical, Verathon Medical, and DEAS Italia; he is a patent co-owner (no royalties) of DEAS Italia and has received lecture grants and travel reimbursements from MSD Italia.

**Name:** Erica Adrario, MD.

**Contribution:** This author helped with substantial manuscript revision.

**Conflicts of Interest:** None.

**Name:** Abele Donati, MD, PhD.

**Contribution:** This author helped with substantial manuscript revision.

**Conflicts of Interest:** None.

**Name:** Stefano Falcetta, MD.

**Contribution:** This author helped conceive the paper, perform literature review, and contribute with substantial manuscript revision.

**Conflicts of Interest:** S. Falcetta has received lecture grants and travel reimbursements from MSD Italia.

**This manuscript was handled by:** Narasimhan Jagannathan, MD, MBA.

## REFERENCES

- Vannucci A, Cavallone LF. Bedside predictors of difficult intubation: a systematic review. *Minerva Anestesiologica*. 2016;82:69–83.
- 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.
- Nørskov AK, Rosenstock CV, Wetterslev J, Astrup G, Afshari A, Lundstrøm LH. Diagnostic accuracy of anaesthesiologists' prediction of difficult airway management in daily clinical practice: a cohort study of 188 064 patients registered in the Danish Anaesthesia Database. *Anaesthesia*. 2015;70:272–281.
- Cattano D, Panicucci E, Paolicchi A, Forfori F, Giunta F, Hagberg C. Risk factors assessment of the difficult airway: an Italian survey of 1956 patients. *Anesth Analg*. 2004;99:1774–1779.
- Levitan RM, Everett WW, Ochroch EA. Limitations of difficult airway prediction in patients intubated in the emergency department. *Ann Emerg Med*. 2004;44:307–313.
- Roth D, Pace NL, Lee A, et al. Bedside tests for predicting difficult airways: an abridged Cochrane diagnostic test accuracy systematic review. *Anaesthesia*. 2019;74:915–928.
- Detsky ME, Jivraj N, Adhikari NK, et al. Will this patient be difficult to intubate?: the rational clinical examination systematic review. *JAMA*. 2019;321:493–503.
- Sorbello M, Falcetta S. Time to include ultrasounds in pre-procedural airway evaluation? *Trends Anaesth Crit Care*. 2021;38:1–3.
- Pandit JJ, Heidegger T. Putting the 'point' back into the ritual: a binary approach to difficult airway prediction. *Anaesthesia*. 2017;72:283–288.
- Peterson GN, Domino KB, Caplan RA, Posner KL, Lee LA, Cheney FW. Management of the difficult airway: a closed claims analysis. *Anesthesiology*. 2005;103:33–39.
- Cook TM, Woodall N, Frerk C; Fourth National Audit Project. Major complications of airway management in the UK: results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: anaesthesia. *Br J Anaesth*. 2011;106:617–631.
- Frerk C, Mitchell VS, McNarry AF, et al; Difficult Airway Society intubation guidelines working group. Difficult Airway Society 2015 guidelines for management of unanticipated difficult intubation in adults. *Br J Anaesth*. 2015;115:827–848.
- De Hert S, Staender S, Fritsch G, et al. Pre-operative evaluation of adults undergoing elective noncardiac surgery: updated guideline from the European Society of Anaesthesiology. *Eur J Anaesthesiol*. 2018;35:407–465.
- Shamseer L, Moher D, Clarke M, et al; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. 2015;350:g7647.
- Salameh JP, Bossuyt PM, McGrath TA, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. *BMJ*. 2020;370:m2632.
- Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia*. 1984;39:1105–1111.
- Yentis SM, Lee DJ. Evaluation of an improved scoring system for the grading of direct laryngoscopy. *Anaesthesia*. 1998;53:1041–1044.
- Reitsma J, Rutjes A, Whiting P, Vlassov V, Leeflang M, Deeks J. Assessing methodological quality. In: Deeks J, Bossuyt P, Gatsonis C, eds. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. 1.0.0. The Cochrane Collaboration, 2009. pp. 1-27. <http://srdta.cochrane.org/>
- Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–536.
- Schünemann HJ, Mustafa RA, Brozek J, et al; GRADE Working Group. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol*. 2020;122:129–141.
- Schünemann HJ, Mustafa RA, Brozek J, et al; GRADE Working Group. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2020;122:142–152.
- Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58:982–990.
- Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20:2865–2884.

24. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 2007;8:239–251.
25. Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res*. 2017;26:1896–1911.
26. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539–1558.
27. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629–634.
28. Ezri T, Gewürtz G, Sessler DI, et al. Prediction of difficult laryngoscopy in obese patients by ultrasound quantification of anterior neck soft tissue. *Anaesthesia*. 2003;58:1111–1114.
29. Komatsu R, Sengupta P, Wadhwa A, et al. Ultrasound quantification of anterior soft tissue thickness fails to predict difficult laryngoscopy in obese patients. *Anaesth Intensive Care*. 2007;35:32–37.
30. Wojtczak JA. Submandibular sonography: assessment of hyomental distances and ratio, tongue size, and floor of the mouth musculature using portable sonography. *J Ultrasound Med*. 2012;31:523–528.
31. Petrisor C, Szabo R, Constantinescu C, Prie A, Hagau N. Ultrasound-based assessment of hyomental distances in neutral, ramped, and maximum hyperextended positions, and derived ratios, for the prediction of difficult airway in the obese population: a pilot diagnostic accuracy study. *Anaesthesiol Intensive Ther*. 2018;50:110–116.
32. Yilmaz C, Karasu D, Dilektasli E, Taha A, Ozgunay SE, Korfali G. An evaluation of ultrasound measurements of anterior neck soft tissue and other predictors of difficult laryngoscopy in morbidly obese patients. *Bariatric Surg Pract Patient Care*. 2018;13:18–24.
33. Sharma A, Bhalla S. Ultrasonographic prediction of difficult laryngoscopy in obese patients. *Indian J Med Spec*. 2020;11:76–80.
34. Pinto J, Cordeiro L, Pereira C, Gama R, Fernandes HL, Assunção J. Predicting difficult laryngoscopy using ultrasound measurement of distance from skin to epiglottis. *J Crit Care*. 2016;33:26–31.
35. Parameswari A, Govind M, Vakamudi M. Correlation between preoperative ultrasonographic airway assessment and laryngoscopic view in adult patients: a prospective study. *J Anaesthesiol Clin Pharmacol*. 2017;33:353–358.
36. Soltani Mohammadi S, Saliminia A, Nejatifard N, Azma R. Usefulness of ultrasound view of larynx in pre-anesthetic airway assessment: a comparison with Cormack-Lehane classification during direct laryngoscopy. *Anesth Pain Med*. 2016;6:e39566.
37. Koundal V, Rana S, Thakur R, Chauhan V, Ekke S, Kumar M. The usefulness of point of care ultrasound (POCUS) in preanaesthetic airway assessment. *Indian J Anaesth*. 2019;63:1022–1028.
38. Abdelhady BS, Elrabiey MA, Elrahman AHA, Mohamed EE. Ultrasonography versus conventional methods (Mallampati score and thyromental distance) for prediction of difficult airway in adult patients. *Egypt J Anaesth*. 2020;36:83–89.
39. Chan SMM, Wong WY, Lam SKT, et al. Use of ultrasound to predict difficult intubation in Chinese population by assessing the ratio of the pre-epiglottis space distance and the distance between epiglottis and vocal folds. *Hong Kong J Emerg Med*. 2018;25:152–159.
40. Falcetta S, Cavallo S, Gabbanelli V, et al. Evaluation of two neck ultrasound measurements as predictors of difficult direct laryngoscopy: a prospective observational study. *Eur J Anaesthesiol*. 2018;35:605–612.
41. Alessandri F, Antenucci G, Piervincenzi E, et al. Ultrasound as a new tool in the assessment of airway difficulties: an observational study. *Eur J Anaesthesiol*. 2019;36:509–515.
42. Senapathi TGA, Wiryana M, Aryabiantara IW, Ryalino C, Roostati RL. The predictive value of skin-to-epiglottis distance to assess difficult intubation in patients who undergo surgery under general anesthesia. *Bali J Anesthesiol*. 2020;4:46–48.
43. Yao W, Wang B. Can tongue thickness measured by ultrasonography predict difficult tracheal intubation? *Br J Anaesth*. 2017;118:601–609.
44. Martínez-García A, Guerrero-Oriach JL, Pino-Gálvez MA. Ultrasonography for predicting a difficult laryngoscopy. Getting closer. *J Clin Monit Comput*. 2021;35:269–277.
45. Ni H, Guan C, He G, Bao Y, Shi D, Zhu Y. Ultrasound measurement of laryngeal structures in the parasagittal plane for the prediction of difficult laryngoscopies in Chinese adults. *BMC Anesthesiol*. 2020;20:134.
46. Wu J, Dong J, Ding Y, Zheng J. Role of anterior neck soft tissue quantifications by ultrasound in predicting difficult laryngoscopy. *Med Sci Monit*. 2014;20:2343–2350.
47. Rana S, Verma V, Bhandari S, Sharma S, Koundal V, Chaudhary SK. Point-of-care ultrasound in the airway assessment: a correlation of ultrasonography-guided parameters to the Cormack-Lehane classification. *Saudi J Anaesth*. 2018;12:292–296.
48. Abo Sabaa MA, Amer GF, Saleh AEAA, Elbakery MAEE. Comparative study between El-Ganzouri airway risk index alone and in combination with upper airway ultrasound in preoperative airway assessment. *Egypt J Hosp Med*. 2019;77:5621–5632.
49. Fulkerson JS, Moore HM, Lowe RF, Anderson TS, Lucas LL, Reed JW. Airway sonography fails to detect difficult laryngoscopy in an adult Veteran surgical population. *Trends Anaesth Crit Care*. 2019;29:26–34.
50. Yadav NK, Rudingwa P, Mishra SK, Pannerselvam S. Ultrasound measurement of anterior neck soft tissue and tongue thickness to predict difficult laryngoscopy - an observational analytical study. *Indian J Anaesth*. 2019;63:629–634.
51. Daggupati H, Maurya I, Singh RD, Ravishankar M. Development of a scoring system for predicting difficult intubation using ultrasonography. *Indian J Anaesth*. 2020;64:187–192.
52. Hui CM, Tsui BC. Sublingual ultrasound as an assessment method for predicting difficult intubation: a pilot study. *Anaesthesia*. 2014;69:314–319.
53. Andruszkiewicz P, Wojtczak J, Sobczyk D, Stach O, Kowalik I. Effectiveness and validity of sonographic upper airway evaluation to predict difficult laryngoscopy. *J Ultrasound Med*. 2016;35:2243–2252.
54. Reddy PB, Punetha P, Chalam KS. Ultrasonography - a viable tool for airway assessment. *Indian J Anaesth*. 2016;60:807–813.
55. Yao W, Zhou Y, Wang B, et al. Can mandibular condylar mobility sonography measurements predict difficult laryngoscopy? *Anesth Analg*. 2017;124:800–806.
56. Wang L, Feng YK, Hong L, et al. Ultrasound for diagnosing new difficult laryngoscopy indicator: a prospective, self-controlled, assessor blinded, observational study. *Chin Med J (Engl)*. 2019;132:2066–2072.
57. Petrisor C, Trancă S, Szabo R, Simon R, Prie A, Bodolea C. Clinical versus ultrasound measurements of hyomental distance ratio for the prediction of difficult airway in patients with and without morbid obesity. *Diagnostics (Basel)*. 2020;10:E140.

58. Shetty SR, Smruthi VT. Validation of clinical versus ultrasound parameters in assessment of airway. *Trends Anaesth Crit Care*. 2020;35:21–27.
59. Adhikari S, Zeger W, Schmier C, et al. Pilot study to determine the utility of point-of-care ultrasound in the assessment of difficult laryngoscopy. *Acad Emerg Med*. 2011;18:754–758.
60. Meco BC, Alanoglu Z, Yilmaz AA, et al. Does ultrasonographic volume of the thyroid gland correlate with difficult intubation? An observational study. *Braz J Anesthesiol*. 2015;65:230–234.
61. Abraham S, Himarani J, Mary Nancy S, Shanmugasundaram S, Krishnakumar Raja VB. Ultrasound as an assessment method in predicting difficult intubation: a prospective clinical study. *J Maxillofac Oral Surg*. 2018;17:563–569.
62. Turkay Aydogmus M, Erkalp K, Nadir Sinikoglu S, Usta TA, O Ulger G, Alagol A. Is ultrasonic investigation of transverse tracheal air shadow diameter reasonable for evaluation of difficult airway in pregnant women: a prospective comparative study. *Pak J Med Sci*. 2014;30:91–95.
63. Xu L, Dai S, Sun L, Shen J, Lv C, Chen X. Evaluation of 2 ultrasonic indicators as predictors of difficult laryngoscopy in pregnant women: a prospective, double blinded study. *Medicine (Baltimore)*. 2020;99:e18305.
64. Reddy A, Aasim S, Satya K, Prasad R. Utility of ultrasonography in preanaesthetic airway assessment. *Asian Pac J Heal Sci*. 2017;4:90–92.
65. Falcetta S, Sorbello M. Bats, Dracula and Batman: the sixth sense in airway management. *Minerva Anesthesiol*. 2021;87:387–390.
66. Cook TM. POGO score. *Can J Anaesth*. 2000;47:477–478.
67. O'Loughlin EJ, Swann AD, English JD, Ramadas R. Accuracy, intra- and inter-rater reliability of three scoring systems for the glottic view at videolaryngoscopy. *Anaesthesia*. 2017;72:835–839.
68. Sorbello M, Frova G. Frova introducer: neither a stylet nor simply an introducer. *Anaesthesia*. 2008;63:1010–1011.
69. Ajithan SE, Puri A, Kapoor MC. Comparison of leakage test and ultrasound imaging to validate ProSeal supraglottic airway device placement. *J Anaesthesiol Clin Pharmacol*. 2020;36:227–232.
70. DI Filippo A, Adembri C, Paparella L, et al; Airway Management Study Group of SIAARTI. Risk factors for difficult Laryngeal Mask Airway LMA-Supreme™ (LMAS) placement in adults: a multicentric prospective observational study in an Italian population. *Minerva Anesthesiol*. 2021;87:533–540.