IMPORTANCE Recognizing patients in whom endotracheal intubation is likely to be difficult can help alert physicians to the need for assistance from a clinician with airway training and having advanced airway management equipment available.

OBJECTIVE To identify risk factors and physical findings that predict difficult intubation.

DATA SOURCES The databases of MEDLINE and EMBASE were searched from 1946 to June 2018 and from 1947 to June 2018, respectively, and the reference lists from the retrieved articles and previous reviews were searched for additional studies.

STUDY SELECTION Sixty-two studies with high (level 1-3) methodological quality that evaluated the accuracy of clinical findings for identifying difficult intubation were reviewed.

DATA EXTRACTION AND SYNTHESIS Two authors independently abstracted data. Bivariate random-effects meta-analyses were used to calculate summary positive likelihood ratios across studies or univariate random-effects models when bivariate models failed to converge.

RESULTS Among the 62 high-quality studies involving 33,559 patients, 10% (95% CI, 8.2%-12%) of patients were difficult to intubate. The physical examination findings that best predicted a difficult intubation included a grade of class 3 on the upper lip bite test (lower incisors cannot extend to reach the upper lip; positive likelihood ratio, 14 [95% CI, 8.9-22]; specificity, 0.96 [95% CI, 0.93-0.97]), shorter hyomental distance (range of <3-5.5 cm; positive likelihood ratio, 6.4 [95% CI, 4.1-10]; specificity, 0.97 [95% CI, 0.94-0.98]), retrognathia (mandible measuring <9 cm from the angle of the jaw to the tip of the chin or subjectively short; positive likelihood ratio, 6.0 [95% CI, 3.1-11]; specificity, 0.98 [95% CI, 0.90-1.0]), and a combination of physical findings based on the Wilson score (positive likelihood ratio, 9.1 [95% CI, 5.1-16]; specificity, 0.95 [95% CI, 0.90-0.98]). The widely used modified Mallampati score (≥3) had a positive likelihood ratio of 4.1 (95% CI, 3.0-5.6; specificity, 0.87 [95% CI, 0.81-0.91]).

CONCLUSIONS AND RELEVANCE Although several simple clinical findings are useful for predicting a higher likelihood of difficult endotracheal intubation, no clinical finding reliably excludes a difficult intubation. An abnormal upper lip bite test, which is easily assessed by clinicians, raises the probability of difficult intubation from 10% to greater than 60% for the average-risk patient.
Clinical Scenario

Case 1
A previously healthy 27-year-old woman was scheduled for elective cholecystectomy. Examination of her airway demonstrated a modified Mallampati score of 2; however, she was unable to bite her upper lip with her lower incisors.

Case 2
A 68-year-old woman with pneumonia was seen on the medical ward for worsening hypoxemia and the need for mechanical ventilation. On initial inspection she was obese, breathing at a respiratory rate of 40 breaths per minute, and had retrognathia. She was confused and uncooperative. Her compromised clinical condition precluded a thorough oropharyngeal and neck examination.

Will endotracheal intubation be difficult in these patients?

Why Is This Question Important?

Endotracheal intubation is often required for major surgical procedures and for respiratory support in critically ill patients. Recognizing a potentially difficult intubation can help clinicians prepare for complications by getting assistance from clinicians with airway training and having advanced airway management equipment available.1,3 Failure to predict and plan for a patient with a difficult airway is the most important factor contributing to the catastrophic “cannot intubate, cannot ventilate” scenario.2-4 Although this occurs in fewer than 1/5000 elective general anesthetic procedures and requires surgical airway rescue in fewer than 1/50 000 cases, these situations can result in major complications associated with long-term morbidity and account for 25% of anesthesia-related deaths.2,4-6 The ability to predict which patients have a high risk of difficult intubation may reduce the risk for “cannot intubate, cannot ventilate” scenarios. This study was performed to identify patient history, clinical features, and bedside tests predictive for difficult intubation.

What Is a Difficult Intubation?

The 2 most common definitions of difficult intubation used in published studies are the Cormack-Lehane grading scale7,8 and the Intubation Difficulty Scale.9 The Cormack-Lehane grading scale describes how visible the vocal cords are during laryngoscopy, ranging from 1 (full view of vocal cords) to 4 (cannot see the epiglottis). The Intubation Difficulty Scale is a scoring system that accounts for the Cormack-Lehane grading scale and other features including the number of intubation attempts, the clinicians involved, advanced airway adjuncts used, the need for increased lifting force, the requirement for external laryngeal pressure, and whether the vocal cords are open or closed during laryngoscopy.

Key Points

Question Which risk factors and physical findings can help predict difficult endotracheal intubation?

Findings In this systematic review, several physical findings increased the likelihood of difficult intubation. The best predictors were an inability to bite the upper lip with the lower incisors, a short hyomental distance, retrognathia, or a combination of findings based on the Wilson score. No risk factor or physical finding consistently ruled out a potentially difficult intubation.

Meaning Although a variety of tests are helpful in identifying a potentially difficult intubation, the inability to bite the upper lip with the lower teeth was the best predictor.

Upper Lip Bite Test, Retrognathia, and Mandibular Protrusion

The upper lip bite test assesses mandibular range of movement by asking the patient to bite their upper lip with their lower incisors. The results of this test are described in terms of 3 grading classifications: class 1, the lower incisors extend beyond the vermilion border of the upper lip; class 2, the lower incisors bite the lip but cannot extend above the vermilion border; and class 3, the lower incisors cannot bite the upper lip at all.15 Among patients without teeth, the upper lip bite test can be replaced with the upper lip catch test, which evaluates whether the lower lip can be raised to cover the vermilion border of the upper lip.16

Retrognathia refers to either the mandible measuring less than 9 cm from the angle of the jaw to the tip of the chin or the subjective appearance of a short mandible. Mandibular protrusion assesses the range of movement of the mandible by asking patients to move their lower teeth past their upper teeth.

Thyromental and Hyomental Distances

The thyromental distance is the distance between the upper-most border of the thyroid cartilage and the mentum measured with the

Components of the Airway Examination

The American Society of Anesthesiologists has identified 11 anatomical features that should be assessed prior to general anesthesia and endotracheal intubation to help identify patients at risk for difficult intubation.10 However, even during emergency situations when a thorough assessment of the oropharynx and neck is not feasible, experienced observers might recognize anthropometric features that increase the likelihood of a difficult intubation. Recognition of the potential for a difficult intubation is the purpose of this review. The factors associated with difficult bag-mask ventilation or establishment of an emergent surgical airway were not reviewed.

History

A comprehensive history begins with a review of prior intubations and factors that may have altered the anatomy of the airway or neck. Examples include previous neck injury, radiation, surgery, or medical conditions including ankylosing spondylitis and diabetes. A history or symptoms suggestive of obstructive sleep apnea should be elicited because this syndrome is associated with upper airway obstruction during sedation.11,12

Physical Examination

Several physical signs and bedside tests have been assessed for predicting difficult endotracheal intubation.13,14 Physical examination should involve inspection of the oropharynx using a penlight and estimates of anthropometric distances and mobility of the cervical spine and mandible.

Upper Lip Bite Test

Retrognathia refers to either the mandible measuring less than 9 cm from the angle of the jaw to the tip of the chin or the subjective appearance of a short mandible. Mandibular protrusion assesses the range of movement of the mandible by asking patients to move their lower teeth past their upper teeth.

Thyromental and Hyomental Distances

The thyromental distance is the distance between the upper-most border of the thyroid cartilage and the mentum measured with the
neck extended. Similarly, the hyomental distance is the distance between the hyoid bone and the mentum (Figure 2). Comparing the thyromental or hyomental distance with a patient’s height can adjust for the difference in these measures in relation to a patient’s overall size. For example, a thyromental distance of 6 cm in a patient who is 200 cm tall is more predictive of difficult intubation than a thyromental distance of 6 cm in a patient who is 160 cm tall. Using a tape measure reduces interobserver variability, but in practice clinicians may use the patient’s fingerbreadths or their own as a surrogate.

Cervical Spine Mobility and Sternomental Distance

The degree of cervical spine flexion and extension as well as any neurological symptoms that arise from neck movement should be assessed prior to intubation. Patients with better cervical spine mobility will have a longer sternomental distance, which is the distance between the upper border of the sternum and the tip of the jaw with the neck fully extended. Poor cervical spine mobility can make intubation more difficult.

Interincisor Gap and Modified Mallampati Score

The maximal distance between the upper and lower incisors is the mouth opening capacity, referred to as the interincisor gap. The modified Mallampati score is a grading system used to rate the visibility of the structures in the oropharynx, including the uvula, faucial pillars, and soft palate when the mouth is opened. The original Mallampati score used a 3-level classification system; however, a modified Mallampati score is more commonly used and has a 4-level system to classify which oropharyngeal structures are visible (Figure 3).

Palm Print Sign and Prayer Sign

Among patients with diabetes, collagen glycosylation can lead to limitations in mobility of the small joints of the hands and other anatomical regions, including the cervical spine. One method to measure mobility of the interphalangeal joints is the palm print sign. An impression of the dominant hand is stamped on a piece of paper and graded based on the proportion of the hand seen on the paper. Another method is the prayer sign, which tests if the patient is able to press his or her 2 palms together.

Abnormal Teeth

Abnormalities in teeth can make it difficult to visualize the vocal cords. This includes subjective assessments of prominent, loose, or missing teeth.

Composite Scores

Combining findings from the history and physical examinations can improve the predictive accuracy for difficult intubation. Composite scores include the El Ganzouri score (which incorporates the modified Mallampati score, interincisor gap, thyromental distance, and cervical spine mobility) and the Wilson Score (which incorporates weight, cervical spine mobility, jaw mobility, degree of retrognathia, and the appearance of the incisors) (Table 1).

Methods

Search Strategy

We conducted a computerized search using OVID versions of MEDLINE (1946-June 2018) and EMBASE Classic and EMBASE (1947-June 2018). The search strategy used was (difficult$ or awkward$ or challenge$ or fail$ or ease or easy or success$ or complicat$ or uncomplicat$) adj2 (intubat$ or airway or laryngoscop$), limited to human. We also searched the reference lists of included studies. Each citation was reviewed in duplicate by 2 of the reviewers, with full-text retrieval of any citation that either reviewer considered potentially relevant for assessing risk factors or clinical tests.
that predict difficult intubation. Additional details appear in eFigure 1 in the Supplement.

**Study Selection**

Two reviewers independently assessed the full text of each retrieved citation. The following criteria were used for study inclusion: (1) cohort study design and a minimum of 10 patients or a clinical trial, (2) population of adult patients aged 18 years or older, (3) intervention of endotracheal intubation performed by direct laryngoscopy, (4) any element of medical history or physical examination, and (5) outcome of difficult laryngoscopy or endotracheal intubation that was measured in the same manner for all patients in each individual study. We excluded studies that were not written in English, were review articles, or if we were unable to abstract relevant data. Studies that used advanced airway devices for endotracheal intubation also were excluded.
Assessment of Study Quality

Study quality was summarized using a quality checklist designed for the Rational Clinical Examination series. Level 1 studies included 100 or more consecutive patients, clinical features were assessed and categorized independently, and the person who intubated the patient was blinded to the assessment. Level 2 studies included less than 100 patients. Level 3 studies included nonconsecutive patients. The study characteristics of level 1 to 3 studies appear in eTable 1 in the Supplement. We excluded level 4 and 5 studies. All studies were graded independently and in duplicate.

Statistical Methods

Two reviewers independently abstracted data to construct 2 × 2 tables for each risk factor and clinical test. Disagreements were arbitrated and resolved by a third reviewer. The 2 × 2 tables were used to calculate sensitivity, specificity, and positive and negative likelihood ratios (LRs). We summarized the sensitivities, specificities, and LRs using a bivariate model when 3 or more studies were available for each topic. When bivariate random-effects models failed to converge, we used a random-effects generic inverse variance method on (1) the logit scale for sensitivity and specificity and (2) the log scale for the LRs. In Table 2, we highlight the results of risk factors and clinical tests that were derived from 3 or more studies and had a summary positive LR of 3 or greater or a summary negative LR less than 0.33 and corresponding 95% CI that exclude 1.0.

When there were only 2 studies for a risk factor or clinical test, the results appear as a range in the Supplement. When the predictive test was described only in a single study, the results
Table 2. Summary of Diagnostic Accuracy for Clinical Tests of Difficult Intubation Evaluated in 3 or More Studies

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Type of Analysis</th>
<th>No. of Patients</th>
<th>No. of Studies</th>
<th>Threshold</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive LR (95% CI)</th>
<th>Negative LR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk factor: snoring[^33-35^]</td>
<td>Bivariate</td>
<td>3866</td>
<td>3</td>
<td>History of snoring</td>
<td>0.43 (0.34-0.53)</td>
<td>0.87 (0.71-0.95)</td>
<td>3.4 (1.6-7.3)</td>
<td>0.65 (0.58-0.72)</td>
</tr>
</tbody>
</table>

**Clinical Tests**

<table>
<thead>
<tr>
<th>Test</th>
<th>Type of Analysis</th>
<th>No. of Patients</th>
<th>No. of Studies</th>
<th>Threshold</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive LR (95% CI)</th>
<th>Negative LR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper lip bite test grading classification</td>
<td>Bivariate</td>
<td>5005</td>
<td>13</td>
<td>Inability of lower incisors to reach lower border of upper lip</td>
<td>0.60 (0.42-0.76)</td>
<td>0.96 (0.93-0.97)</td>
<td>14 (8.9-22)</td>
<td>0.42 (0.27-0.65)</td>
</tr>
<tr>
<td>Class 2 or 3[^33,33,36-50^]</td>
<td>Bivariate</td>
<td>8091</td>
<td>17</td>
<td>Inability of lower incisors to reach upper lip or vermilion border of upper lip</td>
<td>0.60 (0.46-0.73)</td>
<td>0.95 (0.91-0.97)</td>
<td>12 (6.9-20)</td>
<td>0.42 (0.30-0.59)</td>
</tr>
<tr>
<td>Wilson score[^51-58^]</td>
<td>Bivariate</td>
<td>6520</td>
<td>8</td>
<td>≥2 used by 7 studies; ≥3 used by 1 study</td>
<td>0.43 (0.26-0.62)</td>
<td>0.95 (0.90-0.98)</td>
<td>9.1 (5.1-16)</td>
<td>0.60 (0.44-0.82)</td>
</tr>
<tr>
<td>Hyomental distance (measured in neutral position)[^18,42,59^]</td>
<td>Univariate</td>
<td>1245</td>
<td>3</td>
<td>Range, &lt;3-&lt;5.5 cm</td>
<td>0.20 (0.11-0.34)</td>
<td>0.97 (0.94-0.98)</td>
<td>6.4 (4.1-10)</td>
<td>0.84 (0.73-0.96)</td>
</tr>
<tr>
<td>Retrognathia[^25,34,35,40,42^]</td>
<td>Univariate</td>
<td>4017</td>
<td>5</td>
<td>Mandible &lt;9 cm or subjectively short</td>
<td>0.19 (0.07-0.42)</td>
<td>0.98 (0.90-1.0)</td>
<td>6.0 (3.1-11)</td>
<td>0.85 (0.76-0.94)</td>
</tr>
<tr>
<td>Impaired mandibular protrusion[^25,35,53,60-63^]</td>
<td>Bivariate</td>
<td>4229</td>
<td>7</td>
<td>Cannot move lower teeth past upper teeth</td>
<td>0.25 (0.06-0.63)</td>
<td>0.95 (0.86-0.99)</td>
<td>5.5 (2.1-15)</td>
<td>0.78 (0.54-1.1)</td>
</tr>
<tr>
<td>Ratio of height to thyromental distance[^18,40,44,47,48,54^]</td>
<td>Bivariate</td>
<td>3497</td>
<td>6</td>
<td>Range, ≥17-≥25</td>
<td>0.69 (0.57-0.78)</td>
<td>0.87 (0.67-0.95)</td>
<td>5.2 (1.9-14)</td>
<td>0.36 (0.25-0.52)</td>
</tr>
<tr>
<td>Impaired neck mobility[^13,23,26,33,34,40,47,48,48^]</td>
<td>Bivariate</td>
<td>8061</td>
<td>12</td>
<td>Significant variability from subjective assessment to &lt;30°-90° of flexion or extension</td>
<td>0.28 (0.13-0.51)</td>
<td>0.93 (0.85-0.97)</td>
<td>4.2 (1.9-9.5)</td>
<td>0.77 (0.60-0.99)</td>
</tr>
<tr>
<td>Steromental distance[^20,33,40,43,45-47,50,52-54,60,63-66^]</td>
<td>Bivariate</td>
<td>6187</td>
<td>15</td>
<td>Range, ≤2-15 cm</td>
<td>0.41 (0.27-0.57)</td>
<td>0.90 (0.83-0.94)</td>
<td>4.1 (2.7-6.1)</td>
<td>0.65 (0.52-0.82)</td>
</tr>
<tr>
<td>Modified Mallampati score[^15,16,18,23,25,26,33,35-38,41,44,46,47^]</td>
<td>Bivariate</td>
<td>23 396</td>
<td>47</td>
<td>Score ≥3</td>
<td>0.55 (0.48-0.62)</td>
<td>0.87 (0.81-0.91)</td>
<td>4.1 (3.0-5.6)</td>
<td>0.52 (0.45-0.60)</td>
</tr>
<tr>
<td>Impaired mouth opening[^20,33,35,40,43,44,47,50,53,59-61^]</td>
<td>Bivariate</td>
<td>9549</td>
<td>18</td>
<td>Interincisor gap &lt;2-5 cm</td>
<td>0.36 (0.20-0.56)</td>
<td>0.90 (0.80-0.95)</td>
<td>3.6 (2.1-6.1)</td>
<td>0.71 (0.55-0.92)</td>
</tr>
<tr>
<td>Thyromental distance[^23,25,33,40,44,47,50,52-54,60^]</td>
<td>Bivariate</td>
<td>10 596</td>
<td>26</td>
<td>Range, ≤4-&lt;7 cm</td>
<td>0.45 (0.36-0.55)</td>
<td>0.86 (0.80-0.91)</td>
<td>3.3 (2.4-4.4)</td>
<td>0.63 (0.55-0.73)</td>
</tr>
<tr>
<td>Palm print[^12,26,86,87^]</td>
<td>Univariate</td>
<td>695</td>
<td>4</td>
<td>Based on scoring system from each article</td>
<td>0.77 (0.52-0.98)</td>
<td>0.84 (0.55-0.96)</td>
<td>3.0 (1.9-4.7)</td>
<td>0.28 (0.08-0.97)</td>
</tr>
</tbody>
</table>

Abbreviation: LR, likelihood ratio.

[^33-35^]: Measured from the angle of the jaw to the tip of the chin.
appear as a point estimate and 95% CI (eTable 2 in the Supplement). We summarized the pooled incidence of difficult endotracheal intubation on the logit scale using a random-effects generic inverse variance method.

Because standard measures of between-study statistical heterogeneity are not available from bivariate random-effects models, we assessed the consistency of the results using the following sensitivity analyses: (1) restricting the analyses to studies that had a minimum of 20 difficult intubations; (2) excluding higher-risk populations (ie, obstetrical patients, head and neck surgeries, etc) from the analyses; (3) restricting the analyses to studies that used the Cormack-Lehane grading scale as the definition for difficult intubation; and (4) restricting the analyses to studies that fell within first and third quartile of incidence of difficult intubation (ie, 5.7%-15%). For the sensitivity analyses, we calculated summary point estimates and 95% CIs for sensitivities, specificities, and LRIs using the same approach as the primary analysis, but restricted the analyses to predictors that could be summarized using bivariate random-effects models.

When at least 10 studies were available for the same predictor, we evaluated for publication bias that might have favored findings with higher diagnostic accuracy using a weighted regression of the logarithm of the diagnostic odds ratio on the inverse root of the effective sample size.68 We used SAS version 9.4 (SAS Institute Inc) for the bivariate models and R version 3.4.0 (R Foundation for Statistical Computing) for the univariate analysis.

Results

After removal of duplicate studies, the search retrieved 12 394 articles and 62 studies (N = 33 559 patients) met criteria of level 1, 2, or 3 (eTable 1 and eFigure 1 in the Supplement). All studies that were level 1, 2, or 3 were operating room investigations, and some of these studies were restricted to specific patient populations such as obstetric (4 studies),36,51,67,68 patients with diabetes (2 studies),23,86 obese patients (1 study),89 or those undergoing head and neck surgery (3 studies).35,64,73

Incidence of Difficult Intubation

The overall proportion of patients having a difficult intubation was 10% (95% CI, 8.2%-12%). Difficult intubation was most commonly defined as a Cormack-Lehane grade of 3 or 4 (47 studies). Other definitions included the Cormack-Lehane grade with additional requirements (such as the number of intubation attempts, time, or use of bougie; 6 studies), percentage of glottis open (n = 1 study), an Intubation Difficulty Scale score greater than 5 (3 studies), or a minimum intubation time requirement or number of attempts (5 studies) to achieve successful endotracheal intubation.

Risk Factors for Difficult Intubation

A history of difficult intubation (2 studies) was the risk factor most predictive for a difficult intubation (positive LR range, 16-19; negative LR range, 0.72-0.82). Other risk factors included snoring (3 studies; positive LR, 3.4 [95% CI, 1.6-7.3]; negative LR, 0.65 [95% CI, 0.58-0.72]);33-35 difficulty with bag-mask ventilation prior to intubation (1 study; positive LR, 3.5 [95% CI, 2.6-4.7]; negative LR, 0.67 [95% CI, 0.55-0.80]);140 and overweight or obesity (defined as a body mass index >27.35) (5 studies; positive LR, 2.2 [95% CI, 1.6-3.1]; negative LR, 0.70 [95% CI, 0.46-1.1]). Compared with women, men were slightly more difficult to intubate (21 studies; positive LR, 1.2 [95% CI, 1.0-1.3]; negative LR, 0.87 [95% CI, 0.76-0.99]).48,20,26,33,35,37,39,48,52,53,59,62,69,74-76,85 (Table 2 and eTables 2 and 3 in the Supplement).

Accuracy of Clinical Examination

Upper Lip Bite Test, Retrogнатhia, and Mandibular Protrusion

The upper lip bite test (class 3, an inability to bite any part of the upper lip with the lower incisors) strongly predicted a difficult intubation (13 studies; positive LR, 14 [95% CI, 8.9-22]), whereas the ability to extend the teeth above the lower border of the upper lip was predictive of a reduced risk of difficult intubation (negative LR, 0.42 [95% CI, 0.27-0.65]).65,36-44 When including both class 2 and 3 upper lip bite test as a positive test, the results were similar (17 studies; positive LR, 12 [95% CI, 6.9-20]; negative LR, 0.42 [95% CI, 0.30-0.59]).33,36,50 The upper lip catch test, which is used in people with edentulism, had slightly lower predictive accuracy in a single study (positive LR, 7.2 [95% CI, 4.8-11]; negative LR, 0.28 [95% CI, 0.10-0.74]).16

Retrogнатhia (ie, a receding chin [2 studies] or chin length <9 cm [3 studies]) was a good predictor of difficult intubation (positive LR, 6.0 [95% CI, 3.1-11]; negative LR, 0.85 [95% CI, 0.76-0.94]).25,34,35,40,47 Impaired mandibular protrusion (defined as an inability to bring the lower teeth to the upper teeth [2 studies] or past the upper teeth [4 studies]; I study defined it as low protrusion of lower jaw) was also a useful predictor (positive LR, 5.5 [95% CI, 2.1-15]; negative LR, 0.78 [95% CI, 0.54-11]).25,35,53,60-63 (Table 2 and eTable 4 in the Supplement).

Ratio of Height to Thyromental or Hyomental Distance in a Neutral Neck Position vs Neck Extension

A high ratio of height to thyromental distance (6 studies; thresholds ranging from 17 to 25) was predictive of a difficult intubation (positive LR, 5.2 [95% CI, 1.9-14]) and a lower ratio made difficult intubation less likely (negative LR, 0.36 [95% CI, 0.25-0.52]).36,40,44,47,48,54 (Table 2 and eTable 4 in the Supplement). A normal ratio of the hyomental distance measured when the neck is extended compared with when the neck is in a neutral position (1 study; normal is 12) was useful in identifying patients who had an easier intubation (negative LR, 0.19 [95% CI, 0.07-0.56]) (eTables 2 and 4 in the Supplement).

Thyromental and Hyomental Distance

A shorter thyromental distance (thresholds ranging from 4 to 7 cm; 26 studies) increased the likelihood of a difficult intubation (positive LR, 3.3 [95% CI, 2.4-4.4]), whereas a longer thyromental distance made a difficult intubation less likely (negative LR, 0.63 [95% CI, 0.55-0.73]).18,23,25,33,40,43-45,47,50,52-54,60-63,66-69,73,75,76,85 A shorter hyomental distance (thresholds ranging from 3 to 5.5 cm; 3 studies) also was helpful in predicting difficult intubation (positive LR, 6.4 [95% CI, 4.1-10]); negative LR, 0.84 [95% CI, 0.73-0.96]) (Table 2 and eTables 2 and 4 in the Supplement).

Cervical Spine Mobility and Sternotemal Distance

The approach to assessing neck mobility (12 studies) was variable. Definitions included total neck extension of less than 80° (4 studies) or 90° (1 study), atlantaloccipital extension of less than...
35° (2 studies), or other definitions (5 studies). Overall, the presence of impaired neck mobility had modest predictive accuracy (positive LR, 4.2 [95% CI, 1.9-9.5]; negative LR, 0.77 [95% CI, 0.60-0.99]). Sternomental distance (thresholds ranging from <12-15 cm; 15 studies) provided similar results (positive LR, 4.1 [95% CI, 2.7-6.1]; negative LR, 0.65 [95% CI, 0.52-0.82]).

Impaired Mouth Opening
A short interincisor gap (thresholds ranging from <2-5 cm; 18 studies) had moderate accuracy for predicting a difficult intubation (positive LR, 3.6 [95% CI, 2.1-6.1]; negative LR, 0.71 [95% CI, 0.55-0.92]). A modified Mallampati score of 3 or 4 had moderate accuracy for predicting a difficult intubation (positive LR, 4.1 [95% CI, 3.0-5.6]). However, a lower Mallampati score (1 or 2) did not rule out a difficult intubation (negative LR, 0.52 [95% CI, 0.45-0.60]).

Palm Print Sign and Prayer Sign
A positive palm print test result (4 studies) was modestly predictive of a difficult intubation (positive LR, 3.0 [95% CI, 1.9-4.7]), whereas a normal test result made a difficult intubation less likely (negative LR, 0.28 [95% CI, 0.08-0.97]). The prayer sign (defined as no contact between the fourth and fifth metacarpals; 1 study) provided similar results (positive LR, 4.9 [95% CI, 2.8-8.7]; negative LR, 0.75 [95% CI, 0.67-0.84]).

Accuracy of Composite Scores
The Wilson score (8 studies) was the only composite score evaluated in multiple studies in our primary analysis. A Wilson score (≥2 in 7 studies and ≥3 in 1 study) was strongly predictive of a difficult intubation (positive LR, 9.1 [95% CI, 5.1-16]), but a lower score did not exclude difficulty (negative LR, 0.60 [95% CI, 0.44-0.82]). The Lehane grading scale was the most commonly used definition, but overall, the presence of impaired neck mobility had modest predictive accuracy (positive LR, 4.2 [95% CI, 1.9-9.5]; negative LR, 0.77 [95% CI, 0.60-0.99]).

Sensitivity Analyses
For each of the 4 sensitivity analyses of the bivariate results, the point estimates did not qualitatively change the interpretation of the primary results and the 95% CIs tended to be wider given the smaller sample sizes (eFigure 2 in the Supplement).

Publication Bias
For topics with at least 10 studies, there was no evidence of publication bias (ie, suspected unpublished studies with diagnostic odds ratios closer to 1 vs the summary diagnostic odds ratio of published studies) for any of the tests, including sternomental distance (P = .07), impaired mouth opening (P = .71), impaired neck mobility (P = .65), modified Mallampati score (P = .48), sex being male (P = .83), thyromental distance (P = .20), and grade of class 3 on the upper lip bite test (P = .21).

Discussion
An evidence-based approach to predict difficult airway situations should help identify patients who are more likely to be difficult to intubate. Sixty-two high-quality studies were found investigating the accuracy of various risk factors and physical examination findings to predict difficult intubation. The strongest risk factor for difficult intubation is a prior history of difficult intubation; however, the absence of this finding does not rule out difficult intubation. The best bedside test for predicting difficult intubation was the upper lip bite test. Other tests with modest accuracy include hyolaryngeal distance, retrognathia, and impaired mandibular protrusion.

Limitations
First, there was significant variability in the reference standard used among the studies to identify a difficult airway. The Cormack-Lehane grading scale was the most commonly used definition, but it only identifies a difficult view of the vocal cords during direct laryngoscopy rather than a difficult tracheal intubation. Studies that use the number of intubation attempts are vulnerable to differences in clinician ability. Nevertheless, in clinical practice, these definitions are commonly used.

Second, some predictors such as retrognathia and impaired spine mobility require subjective assessments and may be more vulnerable to interobserver variability. There was also significant variation among the studies in how the predictors were defined, thresholds for the various measurements, and in clinician ability.

Third, all level 1 to 3 studies included in this review were conducted in the operating room, which limits applicability to emergency situations. Predictors for difficult intubations in nonemergency situations may still be predictive for emergency situations; however, assessing patients for the risk factors may not be feasible if patients are clinically unstable or unable to follow simple instructions. We restricted our analysis to studies that had independent assessments of predictors and outcomes to minimize bias, but this led to the exclusion of large studies in emergency situations, like the MACHOCA score study.

ETable 4 in the Supplement.

Overall, the presence of impaired neck mobility had modest predictive accuracy (positive LR, 4.2 [95% CI, 1.9-9.5]; negative LR, 0.77 [95% CI, 0.60-0.99]). Sternomental distance (thresholds ranging from <12-15 cm; 15 studies) provided similar results (positive LR, 4.1 [95% CI, 2.7-6.1]; negative LR, 0.65 [95% CI, 0.52-0.82]).

35° (2 studies), or other definitions (5 studies). Overall, the presence of impaired neck mobility had modest predictive accuracy (positive LR, 4.2 [95% CI, 1.9-9.5]; negative LR, 0.77 [95% CI, 0.60-0.99]). Sternomental distance (thresholds ranging from <12-15 cm; 15 studies) provided similar results (positive LR, 4.1 [95% CI, 2.7-6.1]; negative LR, 0.65 [95% CI, 0.52-0.82]).

Impaired Mouth Opening
A short interincisor gap (thresholds ranging from <2-5 cm; 18 studies) had moderate accuracy for predicting a difficult intubation (positive LR, 3.6 [95% CI, 2.1-6.1]; negative LR, 0.71 [95% CI, 0.55-0.92]). A modified Mallampati score of 3 or 4 had moderate accuracy for predicting a difficult intubation (positive LR, 4.1 [95% CI, 3.0-5.6]). However, a lower Mallampati score (1 or 2) did not rule out a difficult intubation (negative LR, 0.52 [95% CI, 0.45-0.60]).

Palm Print Sign and Prayer Sign
A positive palm print test result (4 studies) was modestly predictive of a difficult intubation (positive LR, 3.0 [95% CI, 1.9-4.7]), whereas a normal test result made a difficult intubation less likely (negative LR, 0.28 [95% CI, 0.08-0.97]). The prayer sign (defined as no contact between the fourth and fifth metacarpals; 1 study) provided similar results (positive LR, 4.9 [95% CI, 2.8-8.7]; negative LR, 0.75 [95% CI, 0.67-0.84]).

Accuracy of Composite Scores
The Wilson score (8 studies) was the only composite score evaluated in multiple studies in our primary analysis. A Wilson score (≥2 in 7 studies and ≥3 in 1 study) was strongly predictive of a difficult intubation (positive LR, 9.1 [95% CI, 5.1-16]), but a lower score did not exclude difficulty (negative LR, 0.60 [95% CI, 0.44-0.82]).

Discussion
An evidence-based approach to predict difficult airway situations should help identify patients who are more likely to be difficult to intubate. Sixty-two high-quality studies were found investigating the accuracy of various risk factors and physical examination findings to predict difficult intubation. The strongest risk factor for difficult intubation is a prior history of difficult intubation; however, the absence of this finding does not rule out difficult intubation. The best bedside test for predicting difficult intubation was the upper lip bite test. Other tests with modest accuracy include hyolaryngeal distance, retrognathia, and impaired mandibular protrusion. The Wilson score was the most widely studied composite score and when the score was 2 or greater, it was predictive of a difficult intubation (Table 1). No clinical tests reliably excluded all cases of difficult intubation.

Limitations
First, there was significant variability in the reference standard used among the studies to identify a difficult airway. The Cormack-Lehane grading scale was the most commonly used definition, but it only identifies a difficult view of the vocal cords during direct laryngoscopy rather than a difficult tracheal intubation. Studies that use the number of intubation attempts are vulnerable to differences in clinician ability. Nevertheless, in clinical practice, these definitions are commonly used.

Second, some predictors such as retrognathia and impaired spine mobility require subjective assessments and may be more vulnerable to interobserver variability. There was also significant variation among the studies in how the predictors were defined, thresholds for the various measurements, and in clinician ability.

Third, all level 1 to 3 studies included in this review were conducted in the operating room, which limits applicability to emergency situations. Predictors for difficult intubations in nonemergency situations may still be predictive for emergency situations; however, assessing patients for the risk factors may not be feasible if patients are clinically unstable or unable to follow simple instructions. We restricted our analysis to studies that had independent assessments of predictors and outcomes to minimize bias, but this led to the exclusion of large studies in emergency situations, like the MACHOCA score study.
Fourth, our analysis considered the predictors independently of each other; however, patients may have several factors that increase the risk of difficult intubation.

Fifth, contemporary airway management is less reliant on direct laryngoscopy because there is now greater use of extraglottic airway devices, video laryngoscopy, and advanced airway techniques.10,91,92

### Scenario Resolution

#### Case 1

Reflecting the prevalence of difficult intubation, this patient’s pretest probability of difficult intubation was 10%. Her modified Mallampati score of 2 (negative LR, 0.52) did not suggest she would be difficult to intubate. However, her upper lip bite test grade was class 3 and that grade is associated with a higher likelihood of difficulty (positive LR, 14). The posttest probability of difficulty was 60% based on the upper lip bite test. A video laryngoscope and bougie were made available in the operating room and a second anesthesiologist was present during the intubation attempt. Even though the anesthesiologist’s view of the vocal cords on direct laryngoscopy was a Cormack-Lehane grade of 3, the endotracheal intubation was successful on the first attempt.

#### Case 2

The patient’s pretest probability of difficult intubation was 10%. Based on the cursory physical examination (obese; positive LR, 2.2) and retrognathia (positive LR, 6.0), it was estimated that her posttest probability of a difficult intubation was between 25% and 40%. The patient was transferred to the intensive care unit and a member of the anesthesiology department was called to assist with a plan for the intubation using video laryngoscopy with topical xylocaine and minimal sedation. The patient was intubated successfully on the first attempt with a Cormack-Lehane grade of 2 for the view of the larynx.

#### Clinical Bottom Line

Several individual physical examination findings are predictive but do not reliably exclude the likelihood for a difficult intubation. The most accurate individual bedside clinical assessment is the easily performed upper lip bite test. Given the prevalence of a difficult intubation of 10%, the inability to bite the upper lip with the lower incisors raises the probability of experiencing a difficult intubation to more than 60%. Other individual tests that are helpful include hypomental distance, retrognathia, and impaired mandibular protrusion. The Wilson score is also helpful for predicting which patients will have a difficult intubation.

### ARTICLES INFORMATION

**Accepted for Publication:** December 13, 2018.

**Author Affiliations:** Department of Medicine, Sinai Health System, Toronto, Ontario, Canada (Detsky); Interdepartmental Division of Critical Care Medicine and Department of Medicine, University of Toronto, Toronto, Ontario, Canada (Detsky, Adhikari, Friedrich, Scales); Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada (Jivraj, Adhikari, Wijeysundera, Scales); Department of Anesthesia, University of Toronto, Toronto, Ontario, Canada (Jivraj, Wijeysundera); Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada (Adhikari, Pinto, Scales); Departments of Critical Care Medicine and Medicine, St Michael’s Hospital, Toronto, Ontario, Canada (Friedrich); Li Ka Shing Knowledge Institute, St Michael’s Hospital, Toronto, Ontario, Canada (Friedrich, Wijeysundera); Department of Medicine, Duke University School of Medicine, Durham, North Carolina (Simel); Durham Veterans Affairs Medical Center, Durham, North Carolina (Simel); Department of Anesthesia and Pain Management, University Health Network, Toronto, Ontario, Canada (Wijeysundera).

**Author Contributions:** Drs Detsky and Scales had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

**Concept and design:** Detsky, Adhikari, Friedrich, Wijeysundera, Scales.

**Acquisition, analysis, or interpretation of data:** All authors.

**Drafting of the manuscript:** Detsky, Jivraj, Simel, Scales.

**Critical revision of the manuscript for important intellectual content:** Jivraj, Adhikari, Friedrich, Pinto, Simel, Wijeysundera, Scales.

**Statistical analysis:** Detsky, Pinto, Simel, Scales.

**Administrative, technical, or material support:** Jivraj, Scales.

**Supervision:** Adhikari, Simel, Scales.

**Conflict of Interest Disclosures:** Dr Simel reported receiving honoraria for contributions to JAMA evidence; and is supported by the Durham Center of Innovation to Accelerate Discovery and Practice Transformation (CIN 13-410) at the Durham VA Health Care System. Dr Wijeysundera is supported in part by a New Investigator Award from the Canadian Institutes of Health Research and an Excellence in Research Award from the Department of Anesthesia at the University of Toronto. No other disclosures were reported.

**Additional Contributions:** We acknowledge Jamie Spiegelman, MD (Humber River Hospital, Toronto, Ontario, Canada), for help with a preliminary search and related data abstraction and we thank Daniel Nishijima, MD (University of California-Davis School of Medicine), Karen Welty-Wolf, MD (Durham Veterans Affairs Medical Center and Duke University), and Jonathan Mark, MD (Durham Veterans Affairs Medical Center and Duke University, Durham, NC) for helpful comments on an earlier version of the manuscript. None of those acknowledged were compensated for contributing.

**REFERENCES**


Will This Patient Be Difficult to Intubate?

The Rational Clinical Examination
Clinical Review & Education

Nepal Med Assoc

65 J Anaesthesiol Clin Pharmacol

facio-oro-maxillary and neck malignancy patients.
intubation—a clinical assessment in tests for predicting difficult endotracheal intubation in obese parturients undergoing cesarean delivery.


