

# Reliability of the thyromental height test for prediction of difficult visualisation of the larynx: A prospective external validation

## Address for correspondence:

Dr. Shizuha Yabuki,  
Department of Anesthesiology,  
Tohoku Kosai Hospital,  
2-3-11 Kokubun-cho, Aoba-ku,  
Sendai, 980-0803, Japan.  
E-mail: shizuha.yabuki@gmail.  
com

**Shizuha Yabuki, Satoka Iwaoka, Mamoru Murakami, Hiroko Miura**

Department of Anesthesiology, Tohoku Kosai Hospital, 2-3-11 Kokubun-cho, Aoba-ku, Sendai, Japan

## ABSTRACT

**Background and Aims:** Thyromental height (TMH) has been reported to be useful for prediction of difficult visualisation of the larynx (DVL), defined as Cormack--Lehane (C&L) grade III or IV. The aim of this study was to compare the diagnostic accuracy of the TMH test for DVL with that of other clinically used tests in Japanese patients. **Methods:** Six hundred and nine surgical patients undergoing endotracheal intubation under general anaesthesia were enrolled in this prospective observational study. TMH, thyromental distance (TMD), and Samssoon and Young's modified Mallampati (MMT) tests were performed in all patients. The C&L grades for the laryngoscopic view with and without external backward, upward, rightward pressure (BURP) were determined by designated airway assessors. The cutoff value for the TMH test was calculated using receiver-operating characteristic (ROC) curve analysis. The sensitivity, specificity, positive predictive value, accuracy, positive likelihood ratio, and area under the ROC curve (AUROC) for each predictive test were calculated and compared. **Results:** ROC curve analysis indicated that 54 mm is the optimal cutoff value for the TMH test. However, both this value and the conventional cutoff value of 50 mm, which has been reported as having good diagnostic accuracy in the literature, had poor diagnostic accuracy. The AUROC for the TMH test was 0.631 without BURP and 0.592 with BURP; these values were not superior to those for the TMD test or MMT. **Conclusion:** The TMH test is not a good predictor of DVL in Japanese patients.

**Key words:** Difficult visualisation of the larynx, prospective observational study, thyromental height test

## Access this article online

Website: [www.ijaweb.org](http://www.ijaweb.org)

DOI: 10.4103/ija.IJA\_852\_18

Quick response code



## INTRODUCTION

A difficult airway is one of the complications of airway management and a cause of anaesthesia-related death.<sup>[1]</sup> Accurate assessment and prediction of difficult visualisation of the larynx (DVL), defined as Cormack--Lehane (C&L) grade III or IV, is required to avoid these undesirable outcomes. Although various airway tests have been devised for prediction of DVL in patients with no obvious abnormal anatomy or airway pathology, multiple combinations of single airway tests have been proposed to improve diagnostic accuracy because none of the tests alone has both high sensitivity and high specificity. However, patient-related restrictions and time constraints on the evaluation of multiple factors are commonly encountered, especially in emergencies.

Etezadi *et al.* proposed a new airway prediction test, the thyromental height (TMH) test, which measures the height between the anterior borders of the mentum and the thyroid cartilage with the patient lying supine.<sup>[2]</sup> They found that the TMH test had high sensitivity and high positive predictive value when used alone to predict DVL. A simple reliable test that can predict DVL would be valuable for quick use in

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

**For reprints contact:** [reprints@medknow.com](mailto:reprints@medknow.com)

**How to cite this article:** Yabuki S, Iwaoka S, Murakami M, Miura H. Reliability of the thyromental height test for prediction of difficult visualisation of the larynx: A prospective external validation. *Indian J Anaesth* 2019;63:270-6.

the above-mentioned situations. A new predictive test should undergo external validation before entering clinical use. The reproducibility of the TMH test has been examined but its reliability as a single predictor of DVL remains controversial.<sup>[3-5]</sup>

The aim of this study was to compare the diagnostic accuracy of the TMH test with that of the other tests in clinical use in Japanese patients undergoing general anaesthesia.

## METHODS

This prospective observational study was performed in patients aged  $\geq 18$  years who underwent endotracheal intubation during general anaesthesia between August 2016 and October 2017. Our study protocol was approved by the appropriate institutional review board. Written informed consent was obtained from all patients who participated in the study. The study was registered in a publicly accessible database (UMIN000023219).

The exclusion criteria were as follows: need for awake intubation, emergency surgery, body mass index  $\geq 35$ , the modified Mallampati test (MMT) was too difficult to perform because of, for example, nausea, an obviously difficult airway because of congenital disease or complications, including cervical disc injury, neck injury, and trismus, requirement for rapid sequence intubation, loose teeth that would preclude adequate expansion of the larynx with laryngoscopy, and refusal to give consent.

The following three predictive tests were performed in all patients preoperatively by one of the study investigators, all of whom had more than 3 years of clinical experience in anaesthesia and were appropriately trained to perform the tests. These investigators were not involved in the evaluation of airway difficulty by direct laryngoscopy.

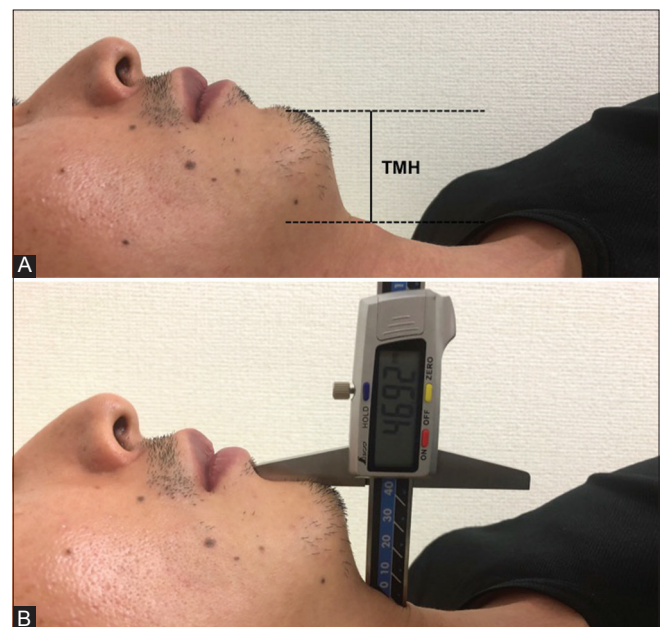
**MMT:** Samssoon and Young's modified MMT was used to assess the visible oropharyngeal structures while the patient was seated with his or her mouth fully opened and the tongue protruded without phonation.<sup>[6]</sup>

**TMD test:** The TMD was measured from the bony point of the mentum to the prominent point of the thyroid cartilage while the patient's head was fully extended and the mouth was closed.<sup>[7]</sup> A TMD  $\leq 65$  mm was considered predictive of DVL.<sup>[8]</sup>

**TMH test:** The TMH was measured from the anterior point of the mentum to the anterior point of the thyroid

cartilage while the patient was lying supine with the mouth closed.<sup>[2]</sup> This height was measured using an electronic depth gauge (Insize<sup>®</sup>, Insize Co Ltd., Suzhou New District, China) [Figure 1].

None of the patients received preanaesthetic medication. For induction of anaesthesia, the patient's head was placed in the sniffing position with a pillow height of 8--10 cm. After adequate preoxygenation, general anaesthesia was induced with intravenous (IV) propofol (1--2 mg/kg), remifentanyl (0.5--1.0  $\mu\text{g}/\text{kg}/\text{min}$ ), and rocuronium (0.6--0.9 mg/kg). Direct laryngoscopy was performed after assessment of neuromuscular blockade 2--5 min after injection of IV rocuronium by one of the airway assessors using a Macintosh #3 or #4 blade with the patient's head tilted. When the laryngoscopic view was estimated to be C&L grade II--IV, the designated airway assessor re-estimated the view when combined with external, backward, upward, rightward pressure (BURP).<sup>[9,10]</sup> The results of the preoperative examinations were not communicated to the investigators who assessed the airway. Easy visualisation of the larynx was defined as C&L grade I or II, and DVL as C&L grade III or IV. All members of the assessment team were anaesthesiology specialists who had qualified in Japan and had more than 10 years' experience performing endotracheal intubation. The preoperative measurements and C&L



**Figure 1:** A Definition of thyromental height: distance from the anterior point of the mentum to the anterior point of the thyroid cartilage while the patient lies supine with the mouth closed. B Measurement of thyromental height with a depth gauge

grades were recorded in a data file on the electronic medical record server.

The cutoff value was calculated and measures of diagnostic accuracy, that is, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and positive likelihood ratio for the TMH test according to C&L grade, were performed with and without BURP. The measures of diagnostic accuracy for the TMH test were then compared with those for the other airway tests.

All statistical analyses were performed using R v. 3.3.1 software (R Foundation for Statistical Computing, Vienna, Austria). The receiver-operating characteristic (ROC) curve was used to calculate the cutoff value for the TMH test. The optimal cutoff point was defined as the point at which the specificity and sensitivity were maximised. The sensitivity, specificity, PPV, NPV, accuracy, and positive likelihood ratio were calculated with and without BURP based on the cutoff value obtained from our own measurements and those described in the literature.<sup>[2]</sup> The area under the ROC curve (AUROC) was calculated for each predictive test and compared using the nonparametric method devised by DeLong *et al.*, which is based on the Mann-Whitney *U* statistic.<sup>[11]</sup> The unpaired Student's *t*-test was used to examine the significance of mean differences in TMH, height, and weight according to patient's gender. A *P* value <0.05 was considered statistically significant in all cases. Considering the objectives of this study and the fact that the rates of DVL in three previous studies of the diagnostic accuracy of the TMH test range from 7.3% to 12%, we assumed that the rate of DVL in our study would be 6%.<sup>[2-4]</sup> A sample size of 609 was deemed necessary to detect agreement between the C&L grade and the TMH test, with a two-sided true Cohen's kappa statistic of 0.6, a kappa under the null hypothesis of 0.4, an alpha level of 5%, and 80% power.

## RESULTS

The demographic data for the 609 patients who participated in the study are shown in Table 1. Five hundred and two patients (82%) were women and the mean age (standard deviation) was 50 (14.1) years. Three hundred and ten patients (50.9%) had undergone breast or gynecological procedures. The ASA physical status was I in 293 patients (48.1%), II in 314 (51.6%), and III in two (0.3%). Laryngoscopy

Variable	Women n=502 (82%)				Men n=107 (18%)			
	Min	Max	Mean	SD	Min	Max	Mean	SD
Age (years)	19	85	49.7	14.1	18	81	52.3	16.7
Height (cm)	136.3	181.5	157.4	5.9	151.3	184.9	169.8	7.0
Weight (kg)	36.0	98.5	56.2	9.6	46.3	98.9	69.7	10.3
BMI	16.0	34.3	22.7	3.6	17.8	32.1	24.1	3.0

BMI – Body mass index; SD – Standard deviation

was difficult in 73 patients (12.0%) without BURP and in six (1.0%) with BURP. Endotracheal intubation was successful in all patients.

When the C&L grade was assessed without BURP, the ROC curve showed that the cutoff value for the TMH test was 54 mm. Previous studies have proposed a cutoff value of 50 mm for the TMH test and 65 mm for the TMD test.<sup>[2,8]</sup> The airway prediction test and C&L grade without BURP are compared in Table 2 and with BURP in Table 3.

Table 4 shows the measures of diagnostic accuracy and AUROC for each test when the C&L grade was assessed without BURP. The sensitivity of the TMH test was 49.3% at a cutoff value of 50 mm and improved to 65.8% at a cutoff value of 54 mm. These sensitivity values were superior to those for the TMD test (6.8%) and MMT (37.0%); however, no value was over 70%. The specificity of the TMD test was 95.7% and its accuracy was 85.1%; these were the highest values of all the tests. The PPV and positive likelihood ratio were highest for the MMT (32.9% and 3.604, respectively). The NPV was highest for the TMH test at a cutoff value of 54 mm (92.3%). The AUROC for the TMH test (0.631) was similar to that for the TMD test (0.636; *P* = 0.911) and the MMT (0.708; *P* = 0.095).

When the C&L grade was assessed with BURP, the sensitivity of the TMH test was 50.0% at a cutoff value of 50 mm and improved to 66.7% at a cutoff value of 54 mm [Table 5]. This value was superior to the sensitivity of the TMD test (16.7%) and the MMT (50.0%). However, the sensitivity did not exceed 70% for any of the tests. The specificity and accuracy were highest for the TMD test (95.5% and 94.7%, respectively). The PPV and positive likelihood ratio were highest for the MMT (3.7% and 3.816, respectively). The NPV was highest for the TMH test at a cutoff value of 54 mm (99.4%). The AUROC for the TMH test (0.592) was similar to that for the TMD test (0.650; *P* = 0.690) but significantly smaller than that for the MMT (0.811; *P* = 0.029).

**Table 2: Comparison of airway prediction tests and Cormack-Lehane grades assessed without BURP**

Prediction Test		C&L Grade		Total	K (95% CI) <sup>b</sup>	Significance (Fisher's exact test)
		Easy (n=536)	Difficult <sup>a</sup> (n=73)			
TMH test ≤50 mm	Easy	378	37	415	0.12 (0.01-0.22)	0.001
	Difficult <sup>c</sup>	158	36	194		
TMH test ≤54 mm	Easy	298	25	323	0.04 (0.00-0.11)	0.001
	Difficult <sup>d</sup>	238	48	286		
TMD test	Easy	513	68	581	0.04 (0.00-0.22)	0.365
	Difficult <sup>e</sup>	23	5	28		
MMT	Easy	481	46	527	0.25 (0.12-0.39)	<0.001
	Difficult <sup>f</sup>	55	27	82		

<sup>a</sup>Grade III and IV views were defined as Difficult on C&L grade; <sup>b</sup>Negative numbers were regarded as zero; <sup>c</sup>≤50 mm was defined as Difficult on TMH test ≤50 mm; <sup>d</sup>≤54 mm was defined as Difficult on TMH test ≤54 mm; <sup>e</sup>≤65 mm was defined as Difficult on TMD test; <sup>f</sup>class III and IV were defined as Difficult on MMT. BURP – External, backward, upward, rightward pressure; CI – Confidence interval; C&L – Cormack-Lehane; MMT – Modified Mallampati test; TMD – Thyromental distance; TMH – Thyromental height

**Table 3: Comparison of airway prediction tests and Cormack-Lehane grade, assessed with BURP**

Prediction Test		C&L Grade		Total	K (95% CI) <sup>b</sup>	Significance (Fisher's exact test)
		Easy (n=603)	Difficult <sup>a</sup> (n=6)			
TMH test ≤50 mm	Easy	412	3	415	0.011 (0.00-0.13)	0.389
	Difficult <sup>c</sup>	191	3	194		
TMH test ≤54 mm	Easy	321	2	323	0.008 (0.00-0.09)	0.427
	Difficult <sup>d</sup>	282	4	286		
TMD test	Easy	576	5	581	0.043 (0.00-0.37)	0.247
	Difficult <sup>e</sup>	27	1	28		
MMT	Easy	524	3	527	0.051 (0.00-0.24)	0.035
	Difficult <sup>f</sup>	79	3	82		

<sup>a</sup>Grade III and IV views were defined as Difficult on C&L Grade; <sup>b</sup>Negative numbers were regarded as zero; <sup>c</sup>≤50 mm was defined as Difficult on TMH test ≤50 mm; <sup>d</sup>≤54 mm was defined as Difficult on TMH test ≤54 mm; <sup>e</sup>≤65 mm was defined as Difficult on TMD test; <sup>f</sup>class III and IV were defined as Difficult on MMT. BURP – External, backward, upward, rightward pressure; CI – Confidence interval; C&L – Cormack-Lehane; MMT – Modified Mallampati test; TMD – Thyromental distance; TMH – Thyromental height

**Table 4: Diagnostic accuracy and AUROC for airway tests predicting C&L grade III or IV assessed without BURP**

Prediction Test	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	Likelihood Ratio (+)	AUROC
TMH test ≤50 mm	49.3 (37.4-61.3) <sup>a</sup>	70.5 (66.5-74.4) <sup>a</sup>	18.6 (13.3-24.8) <sup>a</sup>	91.1 (87.9-93.6) <sup>a</sup>	68.0	1.673	0.631 (0.56-0.70) <sup>a</sup>
TMH test ≤54 mm	<b>65.8</b> (53.7-76.5) <sup>a</sup>	55.6 (51.3-59.9) <sup>a</sup>	16.8 (12.6-21.6) <sup>a</sup>	<b>92.3</b> (88.8-94.9) <sup>a</sup>	56.8	1.481	
TMD test ≤65 mm	6.8 (2.3-15.3) <sup>a</sup>	<b>95.7</b> (93.6-97.3) <sup>a</sup>	17.9 (6.1-36.9) <sup>a</sup>	88.3 (85.4-90.8) <sup>a</sup>	<b>85.1</b>	1.596	0.636 (0.57-0.70) <sup>a</sup>
MMT class III, IV	37.0 (26.0-49.1) <sup>a</sup>	89.7 (86.9-92.2) <sup>a</sup>	<b>32.9</b> (22.9-44.2) <sup>a</sup>	91.3 (88.5-93.5) <sup>a</sup>	83.4	<b>3.604</b>	<b>0.708</b> (0.65-0.77) <sup>a</sup>

AUROC – Area under the receiver-operating characteristic curve; BURP – External, backward, upward, rightward pressure; C&L – Cormack Lehane; MMT – Modified Mallampati test; NPV – Negative predictive value; Likelihood ratio (+) – Positive likelihood ratio; PPV – Positive predictive value; TMD – Thyromental distance; TMH – Thyromental height <sup>a</sup>95% confidence intervals

**Table 5: Diagnostic accuracy and AUROC for airway tests predicting C&L grade III or IV assessed with BURP**

Prediction Test	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	Likelihood Ratio (+)	AUROC
TMH test ≤50 mm	50.0 (11.8-88.2) <sup>a</sup>	68.3 (64.4-72.0) <sup>a</sup>	1.5 (0.3-4.5) <sup>a</sup>	99.3 (97.9-99.9) <sup>a</sup>	68.1	1.579	0.592 (0.32-0.87) <sup>a</sup>
TMH test ≤54 mm	<b>66.7</b> (22.3-95.7) <sup>a</sup>	53.2 (49.2-57.3) <sup>a</sup>	1.4 (0.4-3.5) <sup>a</sup>	<b>99.4</b> (97.8-99.9) <sup>a</sup>	53.4	1.426	
TMD test ≤65 mm	16.7 (0.4-64.1) <sup>a</sup>	<b>95.5</b> (93.6-97.0) <sup>a</sup>	3.6 (0.1-18.3) <sup>a</sup>	99.1 (98.0-99.7) <sup>a</sup>	<b>94.7</b>	3.772	0.650 (0.47-0.83) <sup>a</sup>
MMT class III, IV	50.0 (11.8-88.2) <sup>a</sup>	86.9 (83.9-89.5) <sup>a</sup>	<b>3.7</b> (0.8-10.3) <sup>a</sup>	99.4 (98.3-99.9) <sup>a</sup>	86.5	<b>3.816</b>	<b>0.811</b> (0.70-0.92) <sup>a</sup>

AUROC; area under the receiver-operating characteristic curve; BURP – External, backward, upward, rightward pressure; C&L – Cormack Lehane; MMT – Modified Mallampati test; NPV – Negative predictive value; Likelihood ratio (+) – Positive likelihood ratio; PPV – Positive predictive value; TMD – Thyromental distance; TMH – Thyromental height <sup>a</sup>95% confidence intervals

There was no significant difference in TMH between women and men (56.2 ± 10.6 mm vs. 55.3 ± 11.4 mm; *P* = 0.421); however, there were significant gender related differences in height (157.4 ± 5.9 cm vs. 169.8 cm; *P* < 0.001) and weight (56.2 ± 9.6 kg vs. 69.7 ± 10.3 kg; *P* < 0.001).

We also examined the cutoff value for the TMH test according to gender. The ROC curve for women showed that the cutoff value for the TMH test was 50 mm. In the female population, the sensitivity, specificity, PPV, NPV, accuracy value, and positive likelihood ratio for the TMH test when the C&L grade was

assessed without BURP were 50.0% (95% confidence interval [CI] 36.1--63.9), 70.8% (95% CI 66.3--74.9), 17.1% (95% CI 11.6--23.9), 92.2% (95% CI 88.8--94.8), 68.5%, and 1.71, respectively. The AUROC for the TMH test (0.614, 95% CI 0.53--0.69) was similar to that for the TMD test (0.594, 95% CI 0.51--0.67;  $P = 0.725$ ) and the MMT (0.683, 95% CI 0.61--0.76;  $P = 0.229$ ). When the C&L grade was assessed with BURP, the sensitivity, specificity, PPV, NPV, accuracy value, and positive likelihood ratio for the TMH test were 50.0% (95% CI 1.3--98.7), 68.6% (95% CI 64.3--72.6), 0.6% (95% CI 0.0--0.3), 99.7% (95% CI 98.4--100.0), 68.5%, and 1.59 respectively. The AUROC for the TMH test (0.586, 95% CI 0.01--1.00) was similar to that for the TMD test (0.524, 95% CI 0.37--0.67;  $P = 0.781$ ) and the MMT (0.702, 95% CI 0.67--0.73;  $P = 0.697$ ).

The ROC curve for men showed that the cutoff value for the TMH test was 57 mm. In the male population, the sensitivity, specificity, PPV, NPV, accuracy value, and positive likelihood ratio for the TMH test when the C&L grade was assessed without BURP were 88.9% (95% CI 65.3--98.6), 40.9% (95% CI 30.5--51.9), 23.5% (95% CI 14.1--35.4), 94.7% (95% CI 82.3--99.4), 49.1%, and 1.50, respectively. The AUROC for the TMH test (0.670, 95% CI 0.53--0.81) was similar to that for the TMD test (0.758, 95% CI 0.65--0.87;  $P = 0.197$ ), and the MMT (0.785, 95% CI 0.67--0.90;  $P = 0.128$ ). When the C&L grade was assessed with BURP, the sensitivity, specificity, PPV, NPV, accuracy value, and positive likelihood ratio for the TMH test were 75.0% (95% CI 19.4--99.4), 36.3% (95% CI 27.0--46.4), 4.4% (95% CI 0.9--12.4), 97.4% (95% CI 86.2--99.9), 37.7%, and 1.17, respectively. The AUROC for the TMH test (0.657, 95% CI 0.34--0.97) was similar to that for the TMD test (0.739, 95% CI 0.53--0.95;  $P = 0.694$ ), and the MMT (0.852, 95% CI 0.71--0.99;  $P = 0.0559$ ).

## DISCUSSION

The results of our study indicate that the TMH test is a poor predictor of DVL when used alone. The sensitivity, specificity, PPV, and positive likelihood ratio values for the TMH test were poor for prediction of DVL without BURP [Table 4] and were still poor with BURP [Table 5]. The ROC curve showed that a TMH of 54 mm had maximal sensitivity and specificity, but this new cutoff did not improve each predictive value [Tables 4 and 5]. The low sensitivity of the TMH test suggests that it cannot detect many of the patients in whom laryngoscopy is difficult with or without BURP. Furthermore, the low specificity of the TMH

test indicates that 30%--45% of patients in whom laryngoscopy is straightforward would be misjudged as difficult with or without BURP. Comparison of the AUROCs indicated that the TMH test was not superior to the TMD test or MMT for predicting DVL, irrespective of whether BURP was performed [Tables 4 and 5]. These results indicate that the TMH test is a poor predictor of DVL when used alone.

Whether the TMH test is reliable as a predictor of DVL when used alone has been a matter of debate. An early study by Etezadi *et al.* reported that the TMH test had high diagnostic accuracy in 314 patients (149 men; 47.5%) who underwent general anaesthesia. The sensitivity, specificity, PPV, NPV, and accuracy value were 82.6%, 99.31%, 90.47%, 98.63%, and 98.08%, respectively.<sup>[2]</sup> Jain *et al.* evaluated the diagnostic accuracy of the TMH test in 345 patients undergoing coronary artery bypass grafting under general anaesthesia (gender was not reported) and found that the sensitivity, specificity, PPV, NPV, and accuracy value were 75%, 97%, 73%, 97%, and 95%, respectively.<sup>[3]</sup> Although the sensitivity and PPV in that study were lower than those in the study by Etezadi *et al.*, they were still high. The studies by Etezadi *et al.* and Jain *et al.* indicated that the TMH test could be a predictor of DVL when used alone. In contrast, Selvi *et al.* reported the sensitivity, specificity, PPV, and NPV as 91.89%, 52.17%, 14.66%, and 98.63%, respectively (no accuracy value was reported) in a study of 451 randomly selected patients (230 men; 51.0%) who underwent general anaesthesia.<sup>[4]</sup> The PPV was low in addition to the low specificity. This result indicated that only 14.66% of the cases in which the TMH test predicted DVL would actually be cases of DVL. As mentioned previously, our study showed the sensitivity, specificity, PPV, and positive likelihood ratio for the TMH test to be poor for prediction of DVL [Tables 4 and 5]. Therefore, the latter two studies (by Selvi *et al.* and by our group) indicate that the TMH test is a poor predictor of DVL when used alone.

We believe that the discrepancy between the results of the former two and latter two studies may be attributed to the low incidence of DVL with BURP. Etezadi *et al.* and Jain *et al.* reported incidences of DVL of 7.3% and 9.3%, respectively, as assessed with BURP.<sup>[2,3]</sup> Selvi *et al.* reported that the incidence of DVL was 8.2% without BURP.<sup>[4]</sup> In our study, 12.0% of the patients had DVL without BURP; this number decreased to 1.0% with the addition of BURP. Assessment of the C&L

grade with BURP is considered to be clinically useful because of the direct relationship between DVL with BURP and difficult intubation. If the laryngoscopic view with BURP had been used in the study by Selvi *et al.*, the incidence of DVL would have been lower and the PPV would likely have been poorer. Generally, the PPV is low when the outcome of interest (in this case, DVL) is rare, but it is an important component in a predictive test.<sup>[12]</sup> The low incidence of DVL assessed with BURP in our study (1%) might be a reason for the discrepancy in the results of the four studies. Recently, Rao *et al.* reported sensitivity, specificity, PPV, NPV, and accuracy of 84.62%, 98.97%, 88%, 98.63%, and 97.7%, respectively, in a study of 316 patients (149 men; 47%) who underwent general anaesthesia.<sup>[5]</sup> Although these values are as high as those reported by Etezadi *et al.*, the definition of DVL used by Rao *et al.* was different from that used in the other three studies and in our study.<sup>[2-4]</sup> In their study, DVL was defined as C&L grade IIb and above when assessed without BURP, whereas other studies defined it as C&L grade III or IV. If DVL had been defined as C&L grade III or IV in the study by Rao *et al.*, the incidence of DVL would have decreased from 8.2% (26/316) to 0.9% (3/316). Furthermore, if the laryngoscopic view with BURP had been used in that study, the incidence of DVL would have been much lower and the PPV would likely have been poorer. Therefore, the TMH test might be unreliable, especially in the population with a low incidence of DVL.

Furthermore, like the other predictive tests available, the diagnostic accuracy of the TMH test probably has some heterogeneity, which might account for the poor diagnostic accuracy of this test in our study. The incidence of DVL was low in our study, possibly for the following reasons. First, the patients in this study might simply have had a low rate of DVL. Most of our patients were young women [Table 1] with no severe complications and only two (0.3%) had an ASA physical status score of III. A higher ASA physical status score and male gender have been identified as risk factors for DVL.<sup>[13,14]</sup> Second, the proficiency of the investigators who assessed the airways, all of whom were specialist anaesthesiologists with more than 10 years of experience in endotracheal intubation, could have contributed to the low rate of DVL.<sup>[15,16]</sup> However, for MMT, our sensitivity, specificity, and positive likelihood ratio values were similar to the pooled estimates in two previous meta-analyses (35%, 91%, and 4.1, respectively, and 49%, 86%, and 3.7, respectively).<sup>[17,18]</sup> Furthermore,

for the TMD test, the sensitivity and specificity (20% and 94%, respectively) were similar to the pooled estimates in a previous meta-analysis.<sup>[17]</sup> That is, a low incidence of DVL could not be the only reason for the poor diagnostic accuracy of the TMH test in our study. Shiga *et al.* pointed out the heterogeneity of the sensitivity and specificity of the MMT and the sensitivity of the TMD test in the studies included in their meta-analysis.<sup>[17]</sup> Therefore, the TMH test probably has variable diagnostic accuracy, similar to that of the other predictive tests.

There was no significant difference in TMH between women and men in our study, although gender differences were evident in height and weight. We also examined the cutoff value for the TMH test and the measures of diagnostic accuracy for each gender. However, our study did not reproduce the high diagnostic accuracy for each gender reported by Etezadi *et al.* Furthermore, comparison of the AUROCs indicated that the TMH test was not superior to the TMD test or the MMT for predicting DVL in either gender, irrespective of whether or not BURP was performed. These findings suggest that the gender imbalance resulting from 82% of the patients being female would not have affected the results of the study. Prakash *et al.* reported finding significant gender differences in sternomental distance (SMD), with men having increased morphometric distances.<sup>[19]</sup> They also showed a positive correlation between SMD and patient height, and mentioned the possibility of gender differences affecting the cutoff values for SMD. SMD is the straight distance between the upper border of the manubrium sterni and the bony point of the mentum when the head is fully extended with the mouth closed. Compared with the SMD, the TMH is more focused on the structures of the mandible, including the tongue and soft tissue, and excludes the influence of the length of the neck and the mobility of the head and neck. Although the difference between the SMD and TMH might explain the lack of a significant difference in TMH between the women and men in our study, it is possible that the gender imbalance had affected the results. Therefore, we cannot conclude that there is no gender-related difference in TMH from our present findings alone. Further studies are needed to examine potential gender-related differences in TMH.

This study is the first report on the use of the TMH test in Japan. Thus far, all the studies that have examined the efficacy of the TMH test were conducted in Asia (two in West Asia, two in South Asia, and our study

in East Asia).<sup>[2-5]</sup> There is some research suggesting biological diversity in human cranial traits, even in Asian countries.<sup>[20,21]</sup> Certain cranial traits could determine the value of the TMH as a predictor of DVL. Furthermore, it is possible that the TMH test is simply of limited value, in which case it cannot be recommended for use in East Asia.

Our study has some limitations. First, although we assumed that the rate of DVL would be 6%, the result revealed that the rates were 12% when the C&L grade was assessed without BURP and 1% when the C&L grade was assessed with BURP. Therefore, there is a possibility that the number of study participants was too small, especially when the C&L grade was assessed with BURP. Second, as mentioned previously, there was a predominance of women in our study. However, we examined the significance of differences in TMH between women and men and the cutoff value for the TMH test in each gender to exclude the possibility of a gender bias. Moreover, all our patients were of East-Asian ethnicity. Further studies that include more easily generalisable groups based on race, age, gender, and ASA physical status score may yield more definitive results.

## CONCLUSION

The TMH test alone is not a strong predictor of DVL in Japanese patients. Further research in different races and populations is needed before a definitive conclusion can be reached regarding the ability of this test to predict DVL.

**Financial support and sponsorship**  
Nil.

## Conflicts of interest

There are no conflicts of interest.

## REFERENCES

- Huitink JM, Lie PP, Heideman I, Jansma EP, Greif R, van Schagen N, *et al.* A prospective, cohort evaluation of major and minor airway management complications during routine anaesthetic care at an academic medical centre. *Anaesthesia* 2017;72:42-8.
- Etezadi F, Ahangari A, Shokri H, Najafi A, Khajavi MR, Daghigh M, *et al.* Thyromental height: A new clinical test for prediction of difficult laryngoscopy. *Anesth Analg* 2013;117:1347-51.
- Jain N, Das S, Nanchi M. Thyromental height test for prediction of difficult laryngoscopy in patients undergoing coronary artery bypass graft surgical procedure. *Ann Card Anaesth* 2017;20:207-11.
- Selvi O, Kahraman T, Senturk O, Tulgar S, Serifsoy E, Ozer Z. Evaluation of the reliability of preoperative descriptive airway assessment tests in the prediction of the Cormack-Lehane score: A prospective randomized clinical study. *J Clin Anesth* 2017;36:21-6.
- Rao KVN, Dhatchinamoorthi D, Nandhakumar A, Selvarajan N, Akula HR, Thiruvankatarajan V. Validity of thyromental height test as a predictor of difficult laryngoscopy: A prospective evaluation comparing modified Mallampati score, interincisor gap, thyromental distance, neck circumference, and neck extension. *Indian J Anaesth* 2018;62:603-8.
- Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiberger D, *et al.* A clinical sign to predict difficult laryngoscopy: A prospective study. *Can Anaesth Soc J* 1985;32:429-34.
- Frerk CM. Predicting difficult intubation. *Anaesthesia* 1991;46:1005-8.
- Krobbuaban B, Diregpoke S, Kumkeaw S, Tanomsat M. The predictive value of the height ratio and thyromental distance: Four predictive tests for difficult laryngoscopy. *Anesth Analg* 2005;101:1542-5.
- Knill RL. Difficult laryngoscopy made easy with a "BURP". *Can J Anaesth* 1993;40:279-82.
- Takahata O, Kubota M, Mamiya K, Akama Y, Nozaka T, Matsumoto H, *et al.* The efficacy of the "BURP" maneuver during a difficult laryngoscopy. *Anesth Analg* 1997;84:419-21.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 1988;44:837-45.
- Yentis SM. Predicting difficult intubation-worthwhile exercise or pointless ritual? *Anaesthesia* 2002;57:105-9.
- Riad W, Vaez MN, Raveendran R, Tam AD, Quereshy FA, Chung F, *et al.* Neck circumference as a predictor of difficult intubation and difficult mask ventilation in morbidly obese patients: A prospective observational study. *Eur J Anaesthesiol* 2016;33:244-9.
- Heinrich S, Birkholz T, Irouschek A, Ackermann A, Schmidt J. Incidences and predictors of difficult laryngoscopy in adult patients undergoing general anesthesia: A single-center analysis of 102,305 cases. *J Anesth* 2013;27:815-21.
- Eberhart LH, Arndt C, Gierpka T, Schwaneckamp J, Wulf H, Putzke C. The reliability and validity of the upper lip bite test compared with the Mallampati classification to predict difficult laryngoscopy: An external prospective evaluation. *Anesth Analg* 2005;101:284-9.
- Myneni N, O'Leary AM, Sandison M, Roberts K. Evaluation of the upper lip bite test in predicting difficult laryngoscopy. *J Clin Anesth* 2010;22:174-8.
- Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: A meta-analysis of bedside screening test performance. *Anesthesiology* 2005;103:429-37.
- Lundstrøm LH, Vester-Andersen M, Møller AM, Charuluxananan S, L'hermite J, Wetterslev J; Danish Anaesthesia Database. Poor prognostic value of the modified Mallampati score: A meta-analysis involving 177 088 patients. *Br J Anaesth* 2011;107:659-67.
- Prakash S, Mullick P, Bhandari S, Kumar A, Gogia AR, Singh R. Sternomental distance and sternalmental displacement as predictors of difficult laryngoscopy and intubation in adult patients. *Saudi J Anaesth* 2017;11:273-8.
- Hanihara T, Ishida H, Dodo Y. Characterization of biological diversity through analysis of discrete cranial traits. *Am J Phys Anthropol* 2003;121:241-51.
- Balakrishnan KP, Chockalingam PA. Ethnicity and upper airway measurements: A study in South Indian population. *Indian J Anaesth* 2017;61:622-8.