Original Article

The DIFFMASK score for predicting difficult facemask ventilation: a cohort study of 46,804 patients

L. H. Lundstrøm,[1] **C. V. Rosenstock,**[1] **J. Wetterslev**[2] and **A. K. Nørskov**[3,4]

1 Consultant and Associate Professor, 3 Resident, Department of Anaesthesiology and Intensive Care, Nordsjællands Hospital, Hillerød, Denmark
2 Chief Physician, 4 Resident, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Summary
Facemask ventilation is an essential part of airway management. Correctly predicting difficulties in facemask ventilation may reduce the risk of morbidity and mortality among patients at risk. We aimed to develop and evaluate a weighted risk score for predicting difficult facemask ventilation during anaesthesia. We analysed a cohort of 46,804 adult patients who were assessed pre-operatively airway for 13 predictors of difficult airway management and subsequently underwent facemask ventilation during general anaesthesia. We developed the Difficult Facemask (DIFFMASK) score in two consecutive steps: first, a multivariate regression analysis was performed; and second, the regression coefficients of the adjusted regression model were converted into a clinically applicable weighted point score. The predictive accuracy of the DIFFMASK score was evaluated by assessment of receiver operating characteristic curves. The prevalence of difficult facemask ventilation was 1.06% (95%CI 0.97–1.16). Following conversion of regression coefficients into 0, 1, 2 or 3 points, the cumulated DIFFMASK score ranged from 0 to 18 points and the area under the receiver operating characteristic curve was 0.82. The Youden index indicated a sum score ≥ 5 as an optimal cut-off value for prediction of difficult facemask ventilation giving a sensitivity of 85% and specificity of 59%. The DIFFMASK score indicated that a score of 6–10 points represents a population of patients who may require heightened attention when facemask ventilation is planned, compared with those patients who are obviously at a high- or low risk of difficulties. The DIFFMASK score may be useful in a clinical context but external, prospective validation is needed.

Correspondence to: L. H. Lundstrøm
Email: lars_hyldborg@hotmail.com
Accepted: 24 April 2019
Keywords: airway management; facemask ventilation; pre-operative care
This article is accompanied by an editorial by El-Boghdadly and Aziz, *Anaesthesia* 2019; https://doi.org/10.1111/anae.14703.
Presented in part at the European Anesthesiology Annual Congress, Copenhagen, Denmark, June 2018.

Introduction
Facemask ventilation is an essential part of airway management. Correctly predicting difficulties may allow allocation of adequate resources and potentially reduce the risk of morbidity and mortality among patients at risk [1, 2]. Although pre-operative assessment of every patient's airway is recommended [3–5], several studies have demonstrated a limited ability to predict difficult airway management during general anaesthesia. In a large cohort study, we have shown that the accuracy of clinical prediction of difficult facemask ventilation by anaesthetists in daily clinical practice was poor [6]. Subsequently, we conducted a trial involving 94,000 patients comparing implementation of systematic airway assessment with existing practice on

prediction of difficult facemask ventilation [7]. Twenty-six departments were cluster randomised to assess 11 risk factors for difficult airway management or to continue with their existing airway assessment. Systematic assessment did not alter the overall incidence of unpredicted difficult facemask ventilation; however, there was an increase in the number of false-positive predictions of difficult facemask ventilation. Nevertheless, when the patients who were in fact difficult to facemask ventilate were isolated, the proportion correctly predicted was significantly higher in the intervention group. Even though the prediction of difficult facemask ventilation in the intervention cohort was done after a systematic assessment of risk factors, the protocol did not dictate how this information should be interpreted; that is, when a patient should be categorised as being an 'expected difficult facemask ventilation'. We hypothesised that a weighted score based on an array of risk factors may be an applicable tool for the prediction of difficult facemask ventilation. This study aimed to develop such risk scoring system and then evaluate its diagnostic accuracy for the prediction of difficult facemask ventilation.

Methods

This was a cohort study of patients undergoing general anaesthesia involving facemask ventilation. Data were retrieved from the intervention group of an earlier cluster randomised trial (DIFFICAIR trial [7–10]); the current study, therefore, includes patients from this trial. However, results describing the association between individual risk factors and difficult facemask ventilation and the results of the diagnostic accuracy of the weighted risk score have never been published. This study is reported according to the 'Strengthening the Reporting of Observational Studies in Epidemiology' guidelines [11].

The data were collected in 15 Danish anaesthetic departments between 1 October 2012 and 31 December 2013. A systematic pre-operative registration of risk factors for difficult airway management in adults (age ≥ 15 years) who underwent attempts at facemask ventilation was collated. To help implement and sustain the pre-operative airway assessment, repeated educational sessions regarding the protocol for pre-operative airway assessment were conducted throughout the study period and an array of tutorial aids (video, posters, white-coat aids etc.) was produced and distributed before trial initiation. The Danish Anaesthesia Database served as the platform for data entry. All variables recorded pre-operatively were pre-defined and the registration platform and rules of validation were the same for all patients at each of the registration sites [12].

Facemask ventilation was categorised as follows: (1) easy; (2) difficult – inadequate, unstable or requiring two providers, with or without neuromuscular blocking agents (NMBA); or (3) impossible – unable to facemask ventilate with or without NMBAs. The definitions were a simplification of the four-point grading scale originally proposed by Han and colleagues [13]. In the Danish Anaesthesia Database, grades 1 and 2 are merged into grade 1 (easy), with grades 3 and 4 being represented by grade 2 (difficult) and grade 3 (impossible), respectively. Since previous cohort studies have focused on difficult and impossible facemask ventilation, the results from this cohort study may be directly comparable [14–16]. If not otherwise specified, we analysed patients judged 'impossible' and 'difficult' to ventilate by facemask as a single group.

We assessed all patients using 13 predictors for difficult airway management (Table 1) [7, 8, 14–20]. In addition to the predefined risk factors for difficult facemask ventilation, all anaesthetists had to tick a box stating whether facemask ventilation was anticipated to be difficult or not. The aim was not to influence the anaesthetists to take certain actions for given values of the predefined risk factors associated with difficult airway management. However, they could choose to use these 13 risk factors as guidance in their individual unweighted risk assessment.

In our statistical analyses, the associations between difficult face-mask ventilation and predefined risk factors were assessed by logistic regression and converted to a point score model. The predictive ability of the regression equation and point scores were evaluated by analysis of the area under the receiver operating characteristic (ROC) curve. The prediction models were built in two consecutive steps. First, all the risk factors were included in a multivariate regression analysis, and the final model was derived by backwards elimination at $p < 0.05$. Second, the regression coefficients (β) of the risk factors from the final logistic regression model were converted to a clinically applicable weighted point score. This was done by applying the following rule: ($-0.25 < \beta$ values ≤ 0.25) = 0 point; ($0.25 < \beta$ values ≤ 0.75) = 1 point; ($0.75 < \beta$ values ≤ 1.25) = 2 points; ($1.25 < \beta$ values ≤ 1.75) = 3 points [21].

The diagnostic accuracy of a dichotomised categorisation for each value of the weighted simplified point score [the Difficult Facemask (DIFFMASK) score] was evaluated using the following statistics: sensitivity; specificity; positive predictive value; negative predictive value; positive likelihood ratio; and negative likelihood ratio. The optimal cut-off value of the score was calculated by the Youden index (sensitivity + specificity − 1) [22]. The diagnostic accuracy of

Table 1 List of risk factors for difficult airway management in patients who underwent attempts at facemask ventilation.

Risk factor	Categories	Description of how the risk factor was evaluated
Sex	Female Male	
Age; years	15–44 45–59 60–80 > 80	
Body mass index; kg.m^{-2}	< 25 25–35 ≥ 35	Based on medical records or the patient's own information
Mouth opening; cm	≥ 4 < 4	In patients with incisors the distance between the teeth was measured at maximum mouth opening. In edentulous patients the intergingivale distance was measured at maximum mouth opening
Ability to extend lower jaw	Yes No	The capacity to bring the lower incisors in front of the upper incisors. Edentulous patients are categorised as 'Yes'
Previous difficult tracheal intubation	No Possible Yes, certain	
Thyromental distance; cm	> 6.5 6.0–6.5 < 6.0	The distance was measured along a straight line from the prominentia laryngea of cartilago thyroidea to the notch of mentum mandibulae with maximum head extension
Modified Mallampati score	1/2 3 4	The visibility of the oropharyngeal structures is assessed on the patient sitting in neutral position with maximum mouth opening and tongue protrusion without phonation
Full beard	No Yes	Moustache, goatee or beard stubbles were categorised as 'No'
Snoring	No Yes	
Sleep apnoea	No Yes	History of obstructive sleep apnoea that requires CPAP, BiPAP or surgery
Neck radiation changes	No Yes	
Neck movement; degrees	> 90 80–90 < 80	The range of motion from full extension through full flexion

CPAP, continuous positive airway pressure, BiPAP, bilevel positive airway pressure.

the optimal cut-off value was compared with the previously reported diagnostic accuracy of an unweighted clinical prediction of difficult facemask ventilation by an anaesthetist after performing an airway assessment [7].

The prevalence and pattern of missing data among all covariates was examined. If > 5% of the patients had missing records for one or more covariates, multiple imputations for handling missing data were performed [23–26]. We assumed statistical significance for p value < 0.05 and statistical analysis was performed using SPSS Statistics (version 24.0; IBM, Armonk, NY, USA).

Results

We retrieved records of 46,804 patients with registered risk factors who underwent facemask ventilation. The overall prevalence of difficult/impossible facemask ventilation was 1.06% (95%CI 0.97–1.16). The prevalence of impossible facemask ventilation in isolation was 0.05% (95%CI 0.03–0.07). The characteristics of the patients in relation to the 13 risk factors are shown in Table 2. The proportion of missing data within the various covariates ranged from 0% to 28%; therefore, we performed multiple imputation before further assessment.

Table 2 Characteristics of 46,804 patients undergoing facemask ventilation. Values are number (proportion).

	Difficult facemask ventilation			
	No	**Yes**	**Total**	**Missing**
Sex				–
Female	26,489 (99.3%)	174 (0.7%)	26,663	
Male	19,820 (98.4%)	321 (1.6%)	20,141	
Age; years				–
15–44	16,527 (99.7%)	42 (0.3%)	16,569	
45–59	12,842 (98.9%)	146 (1.1%)	12,988	
60–80	14,465 (98.2%)	272 (1.8%)	14,737	
> 80	2475 (98.6%)	35 (1.4%)	2510	
ASA physical status				–
1–2	42,067 (99.0%)	409 (1.0%)	42,476	
3–5	4242 (98.0%)	86 (2.0%)	4328	
Body mass index; $kg.m^{-2}$				585 (1.2%)
< 25	22,885 (99.5%)	110 (0.5%)	22,995	
25–35	20,769 (98.5%)	314 (1.5%)	21,083	
≥ 35	2073 (96.8%)	68 (3.2%)	2141	
Mouth opening; cm				11,271 (24.1%)
≥ 4	33,361 (98.9%)	372 (1.1%)	33,733	
< 4	1759 (97.7%)	41 (2.3%)	1800	
Ability to extend lower jaw				9841 (21.0%)
Yes	34,406 (98.9%)	370 (1.1%)	34,776	
No	2140 (97.9%)	47 (2.1%)	2187	
Previous difficult tracheal intubation				9406 (20.1%)
No	35,626 (99.0%)	357 (1.0%)	35,983	
Possible	932 (96.6%)	33 (3.4%)	965	
Yes, certain	432 (96.0%)	18 (4.0%)	450	
Thyromental distance; cm				12,042 (25.7%)
> 6.5	29,372 (99.0%)	287 (1.0%)	29,659	
6.0–6.5	4220 (98.0%)	86 (2.0%)	4306	
< 6.0	771 (96.7%)	26 (3.3%)	797	
Modified Mallampati score				8863 (18.9%)
1/2	33,102 (99.1%)	305 (0.9%)	33,407	
3	3799 (97.4%)	100 (2.6%)	3899	
4	608 (95.7%)	27 (4.3%)	635	
Full beard				164 (0.4%)
No	42,864 (99.2%)	354 (0.8%)	43,218	
Yes	3281 (95.9%)	141 (4.1%)	3422	
Snoring				12,776 (27.3%)
No	24,046 (99.3%)	165 (0.7%)	24,211	
Yes	9606 (97.9%)	211 (2.1%)	9 817	
Sleep apnoea				13,032 (27.8%)
No	32,418 (99.0%)	315 (1.0%)	32,733	
Yes	988 (95.1%)	51 (4.9%)	1039	

(continued)

Table 2 (*continued*)

	Difficult facemask ventilation			
	No	Yes	Total	Missing
Neck radiation changes				10,816 (23.1%)
No	35,309 (98.9%)	386 (1.1%)	35695	
Yes	275 (93.9%)	18 (6.1%)	293	
Neck movement; degrees				10,189 (21.8%)
> 90	29,983 (99.1%)	266 (0.9%)	30,249	
80–90	4865 (97.9%)	106 (2.1%)	4971	
< 80	1356 (97.2%)	39 (2.8%)	1395	

All predefined risk factors were significantly associated with difficult facemask ventilation in univariate logistic regression analyses. After including all risk factors in the primary multivariate logistic regression analysis, 'mouth opening', 'neck movement' and 'ability to extend lower jaw' were excluded by stepwise backward elimination as being statistically insignificant covariates when adjusted for other covariates in the final regression model (Table 3). The β values from the various risk factors were converted into points for the DIFFMASK score, resulting in a total score range of 0–18 points. The corresponding cumulated β values ranged from 0 to 9.085 (Table 3).

The diagnostic accuracy of both the DIFFMASK score and cumulated β values had an area under the ROC curve of 0.82 (Fig. 1). The diagnostic accuracy of the DIFFMASK score at various cut-offs is presented in Tables 4 and 5. Equally, the Youden indices indicated either a point score ≥ 5 or ≥ 6 as the optimal cut-off values for predicting difficult facemask ventilation. The prevalence of patients predicted as easy or difficult to facemask ventilate by the unweighted individual risk assessment corresponded to the cut-off values of the DIFFMASK score (Fig. 2). When applying the DIFFMASK score approach, the prevalence of unanticipated difficult facemask ventilation was 15.0% (95%CI 12.0–18.5) ($p < 0.0001$) and 28.9% (95%CI 25.0–33.1%) $p < 0.0001$) for cut-off values of ≥ 5 and ≥ 6, respectively.

Discussion

We found an overall prevalence of difficult/impossible facemask ventilation of 1.1%. In our multivariate regression analysis older age, increased BMI and the presence of a full beard and neck radiation changes were associated with difficult facemask ventilation. We used the independent risk factors from the multivariate regression model to construct the DIFFMASK score for predicting difficult facemask ventilation and an area under the ROC curve of 0.82 indicates that this may be a relatively strong diagnostic test.

Based upon two almost identical Youden indices it was not obvious which cut-off value to choose to discriminate between difficult and easy facemask ventilation. Nevertheless, due to a sensitivity of 85% (and thus a prevalence of unpredicted difficult facemask ventilation of 15%), a cut-off value ≥ 5 points may be optimal because it is crucial to be aware of possible difficult or impossible facemask ventilation pre-operatively. At first sight this appears promising, but at this cut-off value as many as 40% of the patients would erroneously be predicted difficult to ventilate (false positive) and only 2% of the patients having a positive test were actually difficult to ventilate (positive predictive value).

Accurate prediction of difficult facemask ventilation may allow anaesthetists to take precautions to reduce potential complications. In a clinical context, planning airway management according to the positive and negative predictive values of a test may be more relevant than the sensitivity and specificity, as these may be considered in the context of being 'wise after the event'. As in the case of a cut-off value ≥ 5, a poor positive predictive value combined with a high number of false positives may increase the risk of incorrect allocation of resources and create alarm fatigue that paradoxically may cause a safety hazard [27]. This raises the question whether the DIFFMASK score has limited value, and whether this represents another study implying that predicting difficult airway management is a futile ritual [28, 29].

When we compare the performance of the DIFFMASK score with the original unweighted individual risk assessments from the same cohort of patients in the DIFFICAIR trial [7], the positive predictive value was 19% compared to 2% for the DIFFMASK score (≥ 5 points). In the original unweighted risk assessments the fraction of patients predicted difficult to facemask ventilate was < 1% compared to the 41% identified by the DIFFMASK score (≥ 5 points). This demonstrates that the original assessment focused on identifying the patient with the

Table 3 Converting β values of multivariate logistic regression model into points for the DIFFMASK score.

	OR (95%CI)	p value	β value	Point
Age; years				
15–44	Reference			0
45–59	3.15 (2.22–4.46)	< 0.0001	1.147	2
60–80	4.47 (3.21–6.24)	< 0.0001	1.498	3
> 80	5.09 (3.21–8.05)	< 0.0001	1.627	3
Sex				
Female	Reference			0
Male	1.47 (1.18–1.82)	0.001	0.382	1
Body mass index; kg.m^{-2}				
< 25	Reference			0
25–35	2.31 (1.84–5.92)	< 0.0001	0.835	2
> 35	4.24 (3.04–5.92)	< 0.0001	1.445	3
Previous difficult tracheal intubation				
No	Reference			0
Possible	1.51 (1.03–2.23)	0.035	0.415	1
Yes, certain	2.04 (1.16–3.60)	0.014	0.714	1
Thyromental distance; cm				
> 6.5	Reference			0
6.0–6.5	1.36 (1.02–1.82)	0.037	0.311	1
< 6.0	2.30 (1.51–3.49)	< 0.0001	0.831	2
Modified Mallampati score				
1/2	Reference			0
3	1.45 (1.15–1.83)	0.002	0.369	1
4	2.50 (1.53–4.07)	0.001	0.914	2
Full beard				
No	Reference			0
Yes	3.11 (2.48–3.90)	< 0.0001	1.133	2
Snoring				
No	Reference			0
Yes	1.44 (1.17–1.76)	0.001	0.361	1
Sleep apnoea				
No	Reference			0
Yes	1.61 (1.16–2.24)	0.005	0.475	1
Neck radiation changes				
No	Reference			0
Yes	3.33 (1.86–5.96)	< 0.0001	1.203	2
Range of possible accumulated score			0–9.085	0–18

The multivariate logistic regression analysis and the final model were derived by backward elimination: 'mouth opening'; 'neck movement'; and 'ability to extend lower jaw' were all non-significant covariates excluded from the final model.

most obvious risks for difficult facemask ventilation. However, this is not a clinically acceptable approach as it leads to a failure to identify 86% of patients with unpredicted difficultly in facemask ventilation. Consequently, among patients who were truly difficult to facemask ventilate, only 14% were correctly identified (sensitivity). It seems that the original individual unweighted risk assessment approximately corresponded to a cut-off value ranging from ≥ 10 to ≥ 12 points using the DIFFMASK score.

Numerous studies have evaluated individual patient-related risk factors as dichotomous stand-alone tests for

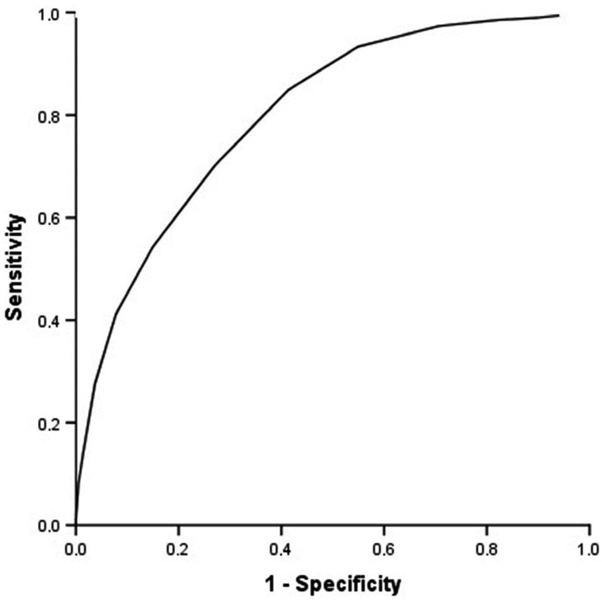

Figure 1 The receiver operator characteristic curve of the DIFFMASK score for predicting difficult facemask ventilation. The area under the curve was estimated to be 0.82 and used as a measure for the description of diagnostic accuracy. The corresponding sensitivity and (1-specificity) of each cut-off value of the DIFFMASK score is presented as the black line.

predicting difficult airway management, but with limited overall diagnostic value [30, 31]. Multivariate sum scores, or the combination of several risk factors, gather more information and potentially increase the diagnostic accuracy [20, 32]. Like stand-alone tests, a multivariate sum score may be used as a dichotomous test with a single cut-off value to discriminate between easy and difficult airway management. In addition, such scores may offer the opportunity to differentiate over a spectrum of increased risk. Our proposed assessment tool does not offer a clear-cut recommendation for a threshold value to discriminate between easy and difficult facemask ventilation. However, a relatively narrow range of DIFFMASK scores (approximately 6–10 points) may be used to pre-operatively identify the patients who may require heightened attention when facemask ventilation is planned, compared with those patients who are obviously at high- or low-risk of difficulties. Thus, if one abandons the assumption that we can rigidly answer whether difficulties with facemask ventilation will occur with a simple ‘yes’ or ‘no’ and instead take certain actions based upon a graduated risk range, then the DIFFMASK score may prove valuable in a clinical context. In practical terms, an online calculator may be used to allow calculation of the DIFFMASK score (www.difficair.com/diffmask-on-line-calculator).

The prevalence of difficult facemask ventilation of 1.1% in the present study is similar to that reported previously (0.9–1.5% [13, 14, 33, 34]). In two publications by Kheterpal et al. involving 50,000 patients undergoing facemask ventilation, six independent risk factors were identified and used to create a prediction score for difficult facemask ventilation [14, 16]. This score dichotomised and included with equal weights all the risk factors, as weighting the factors did not improve the diagnostic accuracy as measured by a ROC curve. In the DIFFMASK score more covariates are incorporated, including more categories which are not dichotomised. As an example, Kheterpal et al.

Table 4 Diagnostic accuracy of the DIFFMASK score for predicting difficult facemask ventilation according to weighted simplified point score cut-off values (Part A). Values are number (proportion).

Cut-off	True positive	False positive	True negative	False negative
≥ 1	489 (1.0%)	41,520 (88.7%)	4789 (10.2%)	6 (0.0%)
≥ 2	488 (1.0%)	38,075 (81.3%)	8234 (17.6%)	7 (0.0%)
≥ 3	482 (1.0%)	32,722 (69.9%)	13,587 (29.0%)	13 (0.0%)
≥ 4	461 (1.0%)	25,430 (54.3%)	20,879 (44.6%)	34 (0.1%)
≥ 5	421 (0.9%)	19,073 (40.8%)	27,236 (58.2%)	74 (0.2%)
≥ 6	352 (0.8%)	12,501 (26.7%)	33,808 (72.2%)	143 (0.3%)
≥ 7	272 (0.6%)	6954 (14.9%)	39,355 (84.1%)	223 (0.5%)
≥ 8	207 (0.4%)	3 632 (7.8%)	42,677 (91.2%)	288 (0.6%)
≥ 9	137 (0.3%)	1702 (3.6%)	44,607 (95.3%)	358 (0.8%)
≥ 10	73 (0.2%)	659 (1.4%)	45,650 (97.5%)	422 (0.9%)
≥ 11	41 (0.1%)	241 (0.5%)	46,068 (98.4%)	454 (1.0%)
≥ 12	19 (0.0%)	90 (0.2%)	46,219 (98.8%)	476 (1.0%)
≥ 13	10 (0.0%)	28 (0.1%)	46,281 (98.9%)	485 (1.0%)

Table 5 Diagnostic accuracy of the DIFFMASK score for predicting difficult facemask ventilation according to weighted simplified point score cut-off values (Part B).

Cut-off	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Youden index	OR (95%CI)
≥ 1	98.8%	10.3%	1.2%	99.9%	1.1	9	9.4 (4.2–21.0)
≥ 2	98.6%	17.8%	1.3%	99.9%	1.2	16	15.1 (7.1–31.8)
≥ 3	97.3%	29.3%	1.5%	99.9%	1.4	27	15.4 (8.9–26.7)
≥ 4	93.1%	45.1%	1.8%	99.8%	1.7	28	11.1 (7.9–15.8)
≥ 5	85.1%	58.8%	2.2%	99.7%	2.1	44[a]	8.1 (6.3–10.4)
≥ 6	71.1%	73.0%	2.7%	99.5%	2.6	44[a]	6.7 (5.5–8.1)
≥ 7	54.9%	85.0%	3.8%	99.4%	3.7	40	6.9 (5.8–8.3)
≥ 8	41.8%	92.2%	5.4%	99.2%	5.3	34	8.4 (7.0–10.1)
≥ 9	27.7%	96.3%	7.5%	99.1%	7.5	24	10.0 (8.2–12.3)
≥ 10	14.7%	98.6%	10.0%	99.1%	10.4	14	12.0 (9.2–15.5)
≥ 11	8.3%	99.5%	14.5%	99.0%	15.9	8	14.4 (10.2–20.2)
≥ 12	3.8%	99.8%	17.4%	99.0%	19.8	4	20.5 (12.4–33.9)
≥ 13	2.0%	99.9%	26.3%	99.0%	33.4	2	34.1 (16.5–70.5)

[a]Optimal cut-off value according to the Youden index (sensitivity + specificity − 100).

dichotomised age by separating groups at an age of 57 years and body mass index (BMI) 30 kg.m^{-2}; we categorised BMI into three categories and age into four categories. Our final model includes 10 covariates, whereas the final model by Kheterpal et al. included six covariates. Furthermore, some of the included covariates differed; for example, jaw protrusion was excluded in our model. Finally, Kheterpal et al. reported the adjusted ORs of the score but not the diagnostic accuracy estimates of sensitivity, specificity, positive and negative predictive values traditionally used.

There are further limitations to our study. Confounding by indication is recognised to introduce bias in non-randomised studies evaluating interventions [35]. The use of NMBAs is thought to improve the conditions for facemask ventilation [36], but we could not

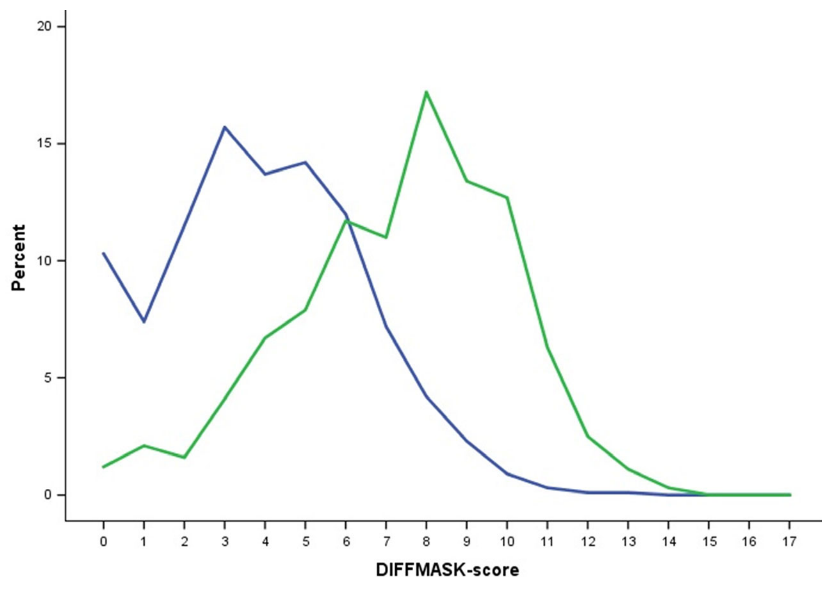

Figure 2 The prevalence of patients either predicted easy (blue line) or difficult (green line) to facemask ventilate by the unweighted individual risk assessment corresponding to the cut-off values of the DIFFMASK scores.

retrieve data on this variable for our assessment. In our study, there was no blinding of the airway assessment and resources may have been allocated to the patients who were obviously at risk of a difficult airway management, thus potentially altering the outcome and diagnostic accuracy of the test. The results could further have been confounded by numerous variables important for the handling (or prediction) of difficult facemask ventilation which were not recorded in our study. For example, it was noticeable that age was identified as a strong independent risk factor. Increasing age may represent other unmeasured patient-related risk factors associated with airway management problems. Among some of the risk factors, the number of missing values was substantial, which may have impacted our findings despite performing multiple imputation [24]. We also did not publish a statistical analysis protocol for data management, analyses and model development, thus our results were explorative to some extent. We suggested the Youden index as a tool for determining the optimal cut-off between easy and difficult facemask ventilation. Although this may be a mathematically correct way to estimate an optimal point, it may not reflect the best cut-off in a clinical setting. As stated, a cut-off value of 5 or 6 points seems problematic, and instead we suggest using the DIFFMASK score as a tool for screening the spectrum of increasing risk of difficulties. The DIFFMASK score was developed and used using the same data for illustrative purposes and, therefore, external validation is needed. Finally, the predictive accuracy of the DIFFMASK score has to be tested and compared with other methods for predicting difficult facemask ventilation.

In conclusion, the DIFFMASK score indicated that a score of 6–10 points represents a population of patients who may require heightened attention when facemask ventilation is planned, compared with those patients who are obviously at a high or low risk of difficulties. However, a completely clear cut-off value separating easy from difficult facemask ventilation was not obvious. The DIFFMASK score may prove valuable in a clinical context but external, prospective validation is needed.

Acknowledgements

The study was funded by independent grants from: TRYG foundation; Nordsjællands Hospitals research council; National Research Council of Denmark; Capital Regions funds for research; and the Danish Society of Anaesthesiology and Intensive Care Medicine. The funding entities had no influence on the study design, conduction or reporting. The research group is independent from the funders. No conflicts of interest declared.

References

1. Cook TM, MacDougall-Davis FR. Complications and failure of airway management. *British Journal of Anaesthesia* 2012; **109**: 68–85.
2. Rosenstock CV, Norskov AK, Wetterslev J, Lundstrom LH. Emergency surgical airway management in Denmark: a cohort study of 452 461 patients registered in the Danish Anaesthesia Database. *British Journal of Anaesthesia* 2016; **117**: 75–82.
3. Apfelbaum JL, Hagberg CA, Caplan RA, et al. Practice guidelines for management of the difficult airway: an updated report by the American Society of Anesthesiologists Task Force on Management of the Difficult Airway. *Anesthesiology* 2013; **118**: 251–70.
4. Cook TM, Woodall N, Frerk C. 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. *British Journal of Anaesthesia* 2011; **106**: 617–31.
5. Frerk C, Mitchell VS, McNarry AF, et al. Difficult Airway Society 2015 guidelines for management of unanticipated difficult intubation in adults. *British Journal of Anaesthesia* 2015; **115**: 827–48.
6. Norskov AK, Rosenstock CV, Wetterslev J, Astrup G, Afshari A, Lundstrom 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**: 272–81.
7. Norskov AK, Wetterslev J, Rosenstock CV, et al. Prediction of difficult mask ventilation using a systematic assessment of risk factors vs. existing practice – a cluster randomised clinical trial in 94,006 patients. *Anaesthesia* 2017; **72**: 296–308.
8. Norskov AK, Wetterslev J, Rosenstock CV, et al. Effects of using the simplified airway risk index vs. usual airway assessment on unanticipated difficult tracheal intubation – a cluster randomized trial with 64,273 participants. *British Journal of Anaesthesia* 2016; **116**: 680–9.
9. Norskov AK, Lundstrom LH, Rosenstock CV, Wetterslev J. Detailed statistical analysis plan for the difficult airway management (DIFFICAIR) trial. *Trials* 2014; **15**: 173.
10. Norskov AK, Rosenstock CV, Wetterslev J, Lundstrom LH. Incidence of unanticipated difficult airway using an objective airway score versus a standard clinical airway assessment: the DIFFICAIR trial – trial protocol for a cluster randomized clinical trial. *Trials* 2013; **14**: 347.
11. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal Clinical Epidemiology* 2008; **61**: 344–9.
12. Antonsen K, Rosenstock CV, Lundstrom LH. The Danish Anaesthesia Database. *Clinical Epidemiology* 2016; **8**: 435–8.
13. Han R, Tremper KK, Kheterpal S, O'Reilly M. Grading scale for mask ventilation. *Anesthesiology* 2004; **101**: 267.
14. Kheterpal S, Han R, Tremper KK, et al. Incidence and predictors of difficult and impossible mask ventilation. *Anesthesiology* 2006; **105**: 885–91.
15. Kheterpal S, Healy D, Aziz MF, et al. Incidence, predictors, and outcome of difficult mask ventilation combined with difficult laryngoscopy: a report from the multicenter perioperative outcomes group. *Anesthesiology* 2013; **119**: 1360–9.
16. Kheterpal S, Martin L, Shanks AM, Tremper KK. Prediction and outcomes of impossible mask ventilation: a review of 50,000 anesthetics. *Anesthesiology* 2009; **110**: 891–7.

17. el-Ganzouri AR, McCarthy RJ, Tuman KJ, Tanck EN, Ivankovich AD. Preoperative airway assessment: predictive value of a multivariate risk index. *Anesthesia and Analgesia* 1996; **82**: 1197–204.
18. Langeron O, Masso E, Huraux C, et al. Prediction of difficult mask ventilation. *Anesthesiology* 2000; **92**: 1229–36.
19. Lundstrom LH, Moller AM, Rosenstock C, Astrup G, Gatke MR, Wetterslev JA. A documented previous difficult tracheal intubation as a prognostic test for a subsequent difficult tracheal intubation in adults. *Anaesthesia* 2009; **64**: 1081–8.
20. 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.
21. Moller MH, Engebjerg MC, Adamsen S, Bendix J, Thomsen RW. The Peptic Ulcer Perforation (PULP) score: a predictor of mortality following peptic ulcer perforation. A cohort study. *Acta Anaesthesiologica Scandinavia* 2012; **56**: 655–62.
22. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; **3**: 32–5.
23. Rubin DB. *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: Wiley online library, 2008.
24. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology* 2017; **17**: 162.
25. Schafer JL. Multiple imputation: a primer. *Statistical Methods in Medical Research* 1999; **8**: 3–15.
26. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychological Methods* 2002; **7**: 147–77.
27. Johnson KR, Hagadorn JI, Sink DW. Alarm safety and alarm fatigue. *Clinics in Perinatology* 2017; **44**: 713–28.
28. Pandit JJ, Heidegger T. Putting the 'point' back into the ritual: a binary approach to difficult airway prediction. *Anaesthesia* 2017; **72**: 283–8.
29. Yentis SM. Predicting difficult intubation–worthwhile exercise or pointless ritual? *Anaesthesia* 2002; **57**: 105–9.
30. Lundstrom LH, Moller AM, Rosenstock C, Astrup G, Wetterslev J. High body mass index is a weak predictor for difficult and failed tracheal intubation: a cohort study of 91,332 consecutive patients scheduled for direct laryngoscopy registered in the Danish Anesthesia Database. *Anesthesiology* 2009; **110**: 266–74.
31. Lundstrom LH, Vester-Andersen M, Moller AM, et al. Poor prognostic value of the modified Mallampati score: a meta-analysis involving 177 088 patients. *British Journal of Anaesthesia* 2011; **107**: 659–67.
32. Cook TM. Strategies for the prevention of airway complications – a narrative review. *Anaesthesia* 2018; **73**: 93–111.
33. Asai T, Koga K, Vaughan RS. Respiratory complications associated with tracheal intubation and extubation. *British Journal of Anaesthesia* 1998; **80**: 767–75.
34. Rose DK, Cohen MM. The airway: problems and predictions in 18,500 patients. *Canadian Journal of Anesthesia* 1994; **41**: 372–83.
35. Deeks JJ, Dinnes J, D'Amico R, et al. Evaluating non-randomised intervention studies. *Health Technology Assessment* 2003; **7**: 1–173.
36. Soltesz S, Alm P, Mathes A, Hellmich M, Hinkelbein J. The effect of neuromuscular blockade on the efficiency of facemask ventilation in patients difficult to facemask ventilate: a prospective trial. *Anaesthesia* 2017; **72**: 1484–90.