e-ISSN: 0975-9506, p-ISSN: 2961-6093

Available online on www.ijpqa.com

International Journal of Pharmaceutical Quality Assurance 2026; 17(9); 380-383

Original Research Article

Agreement between Directly Observed and Standardised Photographic Modified Mallampati Classification: A Prospective Observational Study and Its Implications for Remote Pre-anaesthesia Airway Assessment and Future Artificial Intelligence Development

Gunjan Badwaik[1], Shruti Goswami[2], Anjali Borkar[3], Shweta Takalkar[4]

[1]Professor, Department of Anaesthesia, Datta Meghe Medical College, Nagpur, Maharashtra, India
[2]Professor, Department of Anatomy, Datta Meghe Medical College, Nagpur, Maharashtra, India
[3]Professor and Head, Department of Anaesthesia, Datta Meghe Medical College, Nagpur, Maharashtra, India
[4]Senior Resident, Department of Anaesthesia, Datta Meghe Medical College, Nagpur, Maharashtra, India

Received: 20-09-2026 / Revised: 23-10-2026 / Accepted: 26-10-2026
Corresponding Author: Dr. Gunjan Badwaik
Conflict of interest: Nil

Abstract:

Background: The modified Mallampati classification (MMC) is widely used during pre-anaesthetic airway assessment, but its reproducibility may be affected by posture, phonation, examiner technique and inter-observer variability. Standardised photographic assessment could reduce some sources of variability and may provide a practical bridge to remote airway assessment and future artificial intelligence (AI) applications.

Objective: To evaluate agreement between directly observed MMC performed by a trained anaesthesia junior resident and MMC assigned from standardised oropharyngeal photographs by trained anaesthesia faculty, and to explore whether a standardised photographic dataset could provide a foundation for future AI-assisted remote pre-anaesthesia airway assessment.

Methods: A prospective observational framework involving 2000 adults aged 18–50 years was used. The illustrative dataset contains equal numbers of men and women. A trained JR-2 performed direct MMC assessment, while a trained JR-1 obtained standardised photographs. Anaesthesia faculty independently graded photographs while blinded to the direct score. Agreement was assessed using exact agreement, Cohen's kappa, weighted kappa and symmetry testing.

Results: In the illustrative dataset, exact agreement between direct and photographic MMC was 97.2%, with Cohen's kappa approximately 0.95–0.96, indicating almost-perfect agreement. Most disagreements were between adjacent classes.

Conclusion: Standardised photographic MMC may reproduce direct MMC assessment with high inter-method agreement. Agreement with direct MMC should not be interpreted as proof that photographic MMC predicts difficult intubation. Its principal potential value is standardisation of the human label and creation of a quality-controlled image dataset for future AI-assisted remote airway assessment.

Keywords: Modified Mallampati Classification; Airway Assessment; Photography; Telemedicine; Pre-Anaesthesia Evaluation; Artificial Intelligence; Inter-Rater Reliability; Remote Assessment.

DOI: 10.25258/ijpqa.17.9.57

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0) and the Budapest Open Access Initiative (http://www.budapestopenaccessinitiative.org/read), which permit unrestricted use, distribution, and reproduction in any medium, provided original work is properly credited.

Introduction

Prediction and management of a difficult airway remain central responsibilities of the anaesthesiologist. The Mallampati examination was introduced as a simple visual assessment of oropharyngeal anatomy and was subsequently modified and widely adopted in anaesthetic practice.[1] The test is attractive because it is rapid, inexpensive and familiar, but its performance as a standalone predictor of difficult intubation is limited.[2-5] Reproducibility is another important concern. Mallampati grading can vary with patient posture, phonation, tongue position, examiner instruction, lighting and observer interpretation.[6,7] Meta-analyses and systematic reviews have consequently concluded that individual bedside airway tests have limited sensitivity and that combinations of findings are preferable to reliance on a single test.[3-5,8,9] The growth of telemedicine has introduced a further challenge. Virtual pre-anaesthesia assessment can be feasible and acceptable, but airway examination is more difficult when the anaesthesiologist is not

physically present.[10-12] A standardised photograph could preserve a visual record of the oral examination, allow independent review by several clinicians and reduce some examiner-dependent variation. The digital nature of a photograph also makes it potentially useful for computer-vision research. Recent literature suggests promising applications of AI in airway assessment, but the evidence remains heterogeneous and external validation, dataset quality and clinical integration remain major challenges.[13-17] An important methodological principle is therefore to standardise the label before attempting to automate the label. If human observers do not agree consistently on the classification of an image, an AI system trained on those labels will inherit that uncertainty. The present study was therefore designed to compare directly observed MMC with MMC assessed from standardised photographs and to assess whether such images can provide a reproducible human-labelled dataset suitable for future telemedicine and AI research.

Materials and Methods

Study Design and Setting: This was designed as a prospective observational study conducted over four months at Datta Meghe Medical College, Nagpur. Participants were recruited from outpatient departments of various clinical specialties.

Participants: Adults aged 18–50 years were eligible. The illustrative dataset consisted of 2000 participants, including 1000 men and 1000 women. The final approved protocol should specify all exclusion criteria.

Ethics: Institutional Ethics Committee approval and written informed consent for photography are mandatory. The actual IEC approval number and consent procedure must be inserted before submission.

Direct MMC: A trained anaesthesia JR-2 performed the examination with the patient seated upright, mouth maximally opened and tongue fully protruded without phonation. The modified Mallampati classes were assigned according to standard definitions.[1,6]

Photographic Acquisition: A trained JR-1 obtained a standardised photograph using a predefined protocol covering posture, camera position, distance, lighting, mouth opening and tongue protrusion. Images were de-identified and assigned study codes.

Faculty Assessment: Anaesthesia faculty independently assessed photographs while blinded to the direct MMC score and, ideally, to one another's ratings. A predefined consensus rule should be specified before analysis.

Statistical Analysis: Categorical variables were expressed as frequencies and percentages. Agreement between direct and photographic MMC was evaluated using a 4×4 cross-tabulation, exact percentage agreement, Cohen's kappa and weighted kappa. Symmetry was assessed using the McNemar-Bowker test. Agreement among multiple faculty observers was assessed using Fleiss kappa. Because MMC is ordinal, weighted kappa was considered particularly relevant. A two-sided P value <0.05 was considered statistically significant. The final analysis must be repeated from the raw dataset.

Standardisation Protocol: The patient should be seated upright with the head in a neutral position. The camera should be approximately perpendicular to the facial plane and maintained at a predefined distance. Lighting should permit clear visualisation without glare. The patient should maximally open the mouth and fully protrude the tongue without phonation. Digital zoom should be avoided. Images should be checked immediately against predefined technical-quality criteria. A repeat image should be obtained only when the first image is technically inadequate. This protocol is intended to reduce variability attributable to acquisition rather than to true anatomical differences.

Results — Illustrative/Synthetic Dataset: The following values are simulated for demonstration and are not actual observations.

Table 1:

Direct MMC	Photo I	Photo II	Photo III	Photo IV	Total
I	394	8	0	0	402
II	7	874	15	0	896
III	0	14	525	9	548
IV	0	0	3	151	154
Total	401	896	543	160	2000

The simulated direct MMC distribution was class I 402 (20.1%), class II 896 (44.8%), class III 548 (27.4%) and class IV 154 (7.7%).

Exact agreement was 1944/2000 (97.2%). Disagreement occurred in 56 participants and, in this illustrative matrix, all disagreements were between adjacent grades. The corresponding unweighted Cohen's kappa is approximately 0.95–0.96, indicating almost-perfect agreement. Weighted agreement would be higher because the

disagreement pattern is ordinal and predominantly adjacent.

Discussion

The principal purpose of the present study is not to demonstrate that the Mallampati test is a definitive predictor of difficult intubation. Rather, it addresses whether the same clinical observation can be reproduced from a standardised digital image. This distinction is important because the diagnostic performance of Mallampati alone is modest.[2-5,8,9] If the high agreement demonstrated in the illustrative dataset is reproduced in the actual study, standardised photographs could provide a reproducible visual record of MMC. A second clinician could review the same image, discrepancies could be identified, and the image could be retained for teaching, audit and future research. This may be particularly useful in remote pre-anaesthesia evaluation, where direct positioning and examination by the anaesthesiologist are not possible.[10-12]

The finding would also have implications for AI research. A computer-vision model trained on inconsistent labels will learn label noise. A dataset in which images are acquired under a fixed protocol and labelled independently by several trained anaesthesiologists can provide a more robust starting point. Recent systematic reviews indicate that AI-based airway assessment is promising, but available studies remain heterogeneous and require external validation.[13-17]

Importantly, an AI model trained to reproduce Mallampati class would reproduce a human classification; it would not automatically become a validated predictor of difficult laryngoscopy or intubation. Future studies should therefore link standardised images with clinically meaningful airway outcomes and use multimodal models incorporating relevant clinical variables. External validation across different institutions, devices, age groups, sexes and patient populations will be essential before clinical implementation.[13-17]

For telemedicine, the likely near-term application is clinician-supported rather than fully autonomous. A patient may be guided to acquire a standardised image, an anaesthesiologist may review it remotely, and an AI system could provide a supplementary classification or uncertainty estimate. Abnormal, high-risk or low-confidence results should prompt direct specialist assessment.

Strengths:

The planned sample size of 2000, equal sex representation, predefined age range, separation of image acquisition from direct examination, blinded faculty assessment and use of agreement statistics are important methodological strengths. The photographic dataset may also provide a foundation for future education and AI research.

Limitations: Mallampati classification remains subjective and is an imperfect predictor of difficult airway management.[2-5] This is a single-centre study conducted over four months and may not generalise to all populations. Image quality, camera technology and patient cooperation may affect performance. Most importantly, high agreement with direct MMC does not establish predictive validity for difficult intubation. Finally, all numerical findings in this version are simulated and must be replaced with real data before publication.

Future Research: A multicentre prospective study should link standardised photographs to actual airway outcomes. Future datasets should include diverse populations, multiple camera devices and independent training, validation and external-test cohorts. AI models should be evaluated for calibration, uncertainty, subgroup performance, data leakage and external validity. Ultimately, the clinically relevant endpoint should be difficult airway management rather than photographic reproduction of a Mallampati label.

Conclusion

Standardised photographic assessment of the modified Mallampati classification is a promising method for improving reproducibility and enabling remote review of an examiner-dependent airway observation. In the illustrative dataset, photographic and direct MMC demonstrated almost-perfect agreement. If reproduced using the actual 2000-patient dataset, this would support the use of standardised images as a human-labelled foundation for future telemedicine and AI research. However, photographic MMC should remain one component of comprehensive airway assessment, and agreement with direct MMC should not be interpreted as validation for predicting difficult intubation.

References

1. Samsoon GLT, Young JRB. Difficult tracheal intubation: a retrospective study. Anaesthesia. 1987;42:487-490.
2. Lee A, Fan LTY, Gin T, Karmakar MK, Ngan Kee WD. A systematic review (meta-analysis) of the accuracy of the Mallampati tests to predict the difficult airway. Anesth Analg. 2006;102:1867-1878.
3. Lundstrøm LH, Vester-Andersen M, Møller AM, Charuluxananan S, L'Her E, Wetterslev J. Poor prognostic value of the modified Mallampati score: a meta-analysis involving 177,088 patients. Br J Anaesth. 2011;107:659-667.
4. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently

normal patients: a meta-analysis of bedside screening test performance. Anesthesiology. 2005;103:429-437.
5. Roth D, Pace NL, Lee A, Hovhannisyan K, Warenits AM, Arrich J, Herkner H. Airway physical examination tests for detection of difficult airway management in apparently normal adult patients. Cochrane Database Syst Rev. 2018;5:CD008874.
6. Tham EJ, Gildersleve CD, Sanders LD, Mapleson WW, Vaughan RS. Effects of posture, phonation and observer on Mallampati classification. Br J Anaesth. 1992;68:32-38.
7. Nørskov AK, Rosenstock CV, Wetterslev J, Astrup G, Afshari A, Lundstrøm LH. Inter-observer agreement of tests used for prediction of difficult laryngoscopy/tracheal intubation. Acta Anaesthesiol Scand. 2016;60:101-110.
8. Nørskov AK, Rosenstock CV, Wetterslev J, Astrup G, Afshari A, Lundstrøm LH. Diagnostic accuracy of anaesthesiologists' prediction of difficult intubation in daily clinical practice. Anaesthesia. 2015;70:272-281.
9. JAMA Rational Clinical Examination. Will this patient be difficult to intubate? Systematic review. JAMA. 2019;321:493-503.
10. Azizad S, Joshi GP. Telemedicine for preanesthesia evaluation: review of current evidence. Curr Opin Anaesthesiol. 2021;34:701707.
11. Zhang X, Lee H, et al. Virtual preoperative assessment in surgical patients: a systematic review and meta-analysis. J Clin Anesth. 2022;77:110648.
12. Assessment of the reliability of virtual preanesthetic airway evaluation compared to traditional in-person airway evaluation. J Clin Anesth. 2022.
13. Applications of artificial intelligence in anesthesia: a systematic review. Saudi J Anaesth. 2024.
14. Artificial intelligence in airway management: a systematic review and meta-analysis. Anaesth Crit Care Pain Med. 2025;44(6):101589.
15. Applied artificial intelligence in airway management: a narrative review. Anesthesiol Clin. 2025;43(3):549-561.
16. Artificial intelligence for difficult airway assessment: a protocol for a systematic review with meta-analysis. 2025.
17. Use of artificial intelligence in difficult airway assessment: the current state of knowledge. 2025.
18. Artificial intelligence in airway management. Br J Anaesth. 2026;136(5):1606-1615.
19. AI-based decision models for difficult airway assessment: from research innovation to clinical implementation. 2026.
20. Artificial intelligence in airway management: current evidence and future perspectives. 2025/2026.
21. Frerk C, Mitchell VS, McNarry AF, et al. Difficult Airway Society 2015 guidelines for management of unanticipated difficult intubation in adults. Br J Anaesth. 2015;115:827-848.
22. Apfelbaum JL, Hagberg CA, Connis RT, et al. 2022 American Society of Anesthesiologists practice guidelines for management of the difficult airway. Anesthesiology. 2022;136:31-81.
23. Ahmad I, El-Boghdadly K, Bhagrath R, et al. Difficult Airway Society guidelines for awake tracheal intubation in adults. Anaesthesia. 2020;75:509-528.
24. Cook TM, Woodall N, Frerk C. Major complications of airway management in the UK: results of the Fourth National Audit Project. Br J Anaesth. 2011;106:617-631.
25. Cook TM, Woodall N, Harper J, Benger J. Major complications of airway management in the UK: results of NAP4. Br J Anaesth. 2011;106:632-642.
26. Cook TM, MacDougall-Davis SR. Complications and failure of airway management. Br J Anaesth. 2012;109 Suppl 1:i68-i85.
27. A comparative study to evaluate the efficacy of virtual versus direct airway assessment. J Anaesthesiol Clin Pharmacol. 2021.
28. Remote, automatic, digital preanesthetic evaluation—are we there yet? J Clin Anesth. 2024.
29. Artificial intelligence and telemedicine in anesthesia: potential and problems. J Clin Anesth. 2022.
30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370:1453-1457.