Check for updates

Review Article

Videolaryngoscopy vs. direct Macintosh laryngoscopy in tracheal intubation in adults: a ranking systematic review and network meta-analysis

C. C. de Carvalho,[1] **D. M. da Silva,**[2] **V. M. Lemos,**[3] **T. G. B. dos Santos,**[4] **I. C. Agra,**[5] **G. M. Pinto,**[6] **I. B. Ramos,**[7] **Y. S. C. Costa**[8] and **J. M. Santos Neto**[9]

1 Professor, Department of Surgery, Universidade Federal de Campina Grande, Campina Grande, Brazil
2 Head of Department, 3,4,9 Attending Anaesthetist, Support and Therapeutic Diagnosis Division, Anaesthesiology and Post-Anaesthetic Care Unit, Hospital das Clínicas da Universidade Federal de Pernambuco, Recife, Brazil
5,6,7,8 Student, Centro de Ciências Biológicas e da Saúde, Universidade Federal de Campina Grande, Campina Grande, Brazil

Summary

Videolaryngoscopes are thought to improve glottic view and facilitate tracheal intubation compared with the Macintosh direct laryngoscope. However, we currently do not know which one would be the best choice in most patients undergoing anaesthesia. We designed this systematic review with network meta-analyses to rank the different videolaryngoscopes and the Macintosh direct laryngoscope. We conducted searches in PubMed and a further five databases on 11 January 2021. We included randomised clinical trials with patients aged ≥16 years, comparing different videolaryngoscopes, or videolaryngoscopes with the Macintosh direct laryngoscope for the outcomes: failed intubation; failed first intubation attempt; failed intubation within two attempts; difficult intubation; percentage of glottic opening seen; difficult laryngoscopy; and time needed for intubation. We assessed the quality of evidence according to GRADE recommendations and included 179 studies in the meta-analyses. The C-MAC and C-MAC D-Blade were top ranked for avoiding failed intubation, but we did not find statistically significant differences between any two distinct videolaryngoscopes for this outcome. Further, the C-MAC D-Blade performed significantly better than the C-MAC Macintosh blade for difficult laryngoscopy. We found statistically significant differences between the laryngoscopes for time to intubation, but these differences were not considered clinically relevant. The evidence was judged as of low or very low quality overall. In conclusion, different videolaryngoscopes have differential intubation performance and some may be currently preferred among the available devices. Furthermore, videolaryngoscopes and the Macintosh direct laryngoscope may be considered clinically equivalent for the time taken for tracheal intubation. However, despite the rankings from our analyses, the current available evidence is not sufficient to ensure significant superiority of one device or a small set of them over the others for our intubation-related outcomes.

Correspondence to: C. de Carvalho
Email: clistenescristian@hotmail.com
Accepted: 14 October 2021
Keywords: airway management; laryngoscopy; network meta-analysis; systematic review; tracheal intubation

Introduction

Difficult airways are important causes of major complications during peri-operative care [1]. However, we are currently not able to accurately predict the occurrence of difficult intubation [2–5]. Videolaryngoscopes may offer some benefit [2, 6–9]; several studies have shown the potential

© 2021 Association of Anaesthetists

use of these devices for tracheal intubation in comparison with direct laryngoscopy [8, 10–14]. Additionally, available evidence has also highlighted possible differential intubating performances among the different videolaryngoscopes [8, 9]. However, it is not clear which device we should choose to intubate patients from different populations across different scenarios [8, 9]. We therefore designed this systematic review with network meta-analyses to rank videolaryngoscopes for orotracheal intubation performance compared with the Macintosh direct laryngoscope in adult patients.

Methods

Our inclusion criteria were as follows: randomised clinical trials fully reported; human patients from any population (e.g. elderly patients, neck immobilisation, pregnant women, obese patients, general population, etc.) aged ≥16 y; data available on failed intubation with the device, failed first intubation attempt, number of intubation attempts, difficulty of intubation, percentage of glottic opening, Cormack and Lehane classification and time for intubation; and comparison between videolaryngoscopes or between videolaryngoscopes and the Macintosh direct laryngoscope. We excluded: studies published in languages other than English, Spanish or Portuguese; studies where it was impossible to abstract relevant data on outcomes – including contradictory data; and studies with systematic differences in the intubation technique, including drugs used, between the intervention groups other than the laryngoscopes.

We conducted a computerised search through PubMed, LILACS, SciELO, Embase, Web of Science and Cochrane Central Register of Controlled Trials (CENTRAL) on 11 January 2021. We also searched the reference lists of the included studies. The following search strategy line was applied to PubMed with no limitations: "(laryngoscopes [MeSH] OR laryngoscop* OR videolaryngoscop* OR GlideScope OR Pentax OR C-MAC OR blade OR McGrath OR X-lite OR Airtraq OR Trueview OR CEL-100 OR "King vision" OR Bullard OR Venner OR vividtrac OR "copilot VL" OR "ue?scope") AND ("Airway management" [MeSH] OR "Airway management" OR intubation* OR difficult* OR visualization OR view)". Similar search strategies were applied to the other databases.

The retrieved references were taken into the EPPI Reviewer Web (Beta) for the screening steps – 'title and abstract' then 'full text' [15]. The eligibility criteria were applied to select the studies to be included. Four pairs of reviewers (JSN and YC; DS and GP; VL and IR; and TS and IA) performed in duplicate and independently all steps from screening of title and abstract, through screening of full text and risk of bias assessment to data extraction. The results were compared with each other and disagreements solved by discussion and consensus among the correspondent researchers and the first author (CC). If agreement could not be reached, CC acted as the final judge. We tried to reach authors when information was missing and excluded those outcomes where relevant data were not presented or were conflicting to one another, and the corresponding authors did not reply to our contact after three attempts over a period of 1 month. Data were recorded in Microsoft Excel™ spreadsheets. The data extraction tools were first tested in five included studies and then refined if necessary. We handled the data from each study group (each device) separately for network analyses. For studies including the Macintosh direct laryngoscope, we also combined data from videolaryngoscopes' groups into a single group from multi-group studies to create a pairwise comparison for the pairwise analyses (videolaryngoscopes vs. Macintosh direct laryngoscope).

We collected or calculated data using a standardised tool to include: author's name; publication year; study design; patient characteristics; mean age; mean BMI; mean weight; mean height; sex; ASA physical status; setting; country; sampling; nature of procedure (elective vs. urgent); intubation technique (standard vs. rapid sequence induction); operator experience; number of participants randomised and analysed; number of participants in each group; type of laryngoscope; induction agent and dose; opioid used and dose; and neuromuscular blocking drug used and dose. We defined an experienced operator as one who had performed at least 20 intubations with the device. The use of neuromuscular blocking drugs and opioids was considered as separate variables and not as integral to the intubation techniques.

Our primary outcome was the risk of failed intubation with the devices. Our secondary outcomes were: failed first intubation attempt; failed intubation within two attempts; difficulty of intubation; percentage of glottic opening seen; difficult laryngoscopy; and time to intubation. We standardised the outcomes' events as negative events (the higher the relative risk or the mean difference, the worse the device performance), except for percentage of glottic opening, whose event was taken as positive (the higher the mean difference in percentage of glottic opening, the better the intervention performance).

We built a network graph to evaluate the overall arrangement of the network evidence base. A network graph is made up of nodes and lines. The nodes depict what we regarded as individual interventions (laryngoscopes), whereas the lines connecting different nodes represent the direct comparisons between the relevant devices, and their

© 2021 Association of Anaesthetists

thicknesses are proportional to the number of randomised trials that studied the respective direct comparison.

We applied the Cochrane Risk of Bias 2 tool to assess the risk of bias of the individual studies for each outcome [16]. Five domains are assessed through this tool: randomisation process; deviation from intended interventions; missing outcome data; measurement of the outcome; and selection of reported results. An overall risk of bias assessment was also performed.

We conducted both pairwise and network meta-analyses. The former considered videolaryngoscopes vs. Macintosh direct laryngoscope, whereas the latter considered individual devices. Data were summarised in the pairwise analyses if at least two different sources were available. The analyses were conducted using Review Manager v5.3.5 (RevMan, London, UK) and R software tools (R Foundation for Statistical Computing, Vienna, Austria), as appropriate. The dataset, as well as the analytical code are available at Mendeley Data (https://data.mendeley.com/datasets/bpkc8k9wnx/1) and https://rpubs.com/clistenescarvalho/Anaesthesia, respectively. Per-protocol raw outcome data (i.e. not pre-calculated effect size data) were extracted or calculated from studies and summarised. Effect sizes, standard errors and 95%CI were estimated for each study from the recorded data. For the pairwise analyses, forest plots of relative risk or standardised mean difference were constructed for every outcome. Pooled estimates were calculated by fixed-effects (Mantel–Haenszel or inverse variance method, where appropriate) and random-effects (Sidik-Jonkman method with Hartung-Knapp adjustment) for sensitivity analyses. Heterogeneity was evaluated qualitatively and quantitatively by Cochran's Q test and I^2. Where qualitative or quantitative heterogeneity was present, pooled estimates from random-effects models were presented. An influence analysis by the 'leave-one-out' method was performed to assess the influence of each study on the pooled effects and the heterogeneity between studies. Additionally, to deal with the risk of type-1 and type-2 errors due to repeated significance testing by subsequent meta-analyses, we applied trial sequential analysis to the main outcome.

We performed Bayesian random-effects network meta-analyses using the R package 'gemtc' [17]. These analyses concomitantly considered the devices specified in the included studies for direct and indirect comparisons. We divided out multi-group studies into multiple pairwise comparisons, taking one single group (one device) as the reference to which all the other groups were compared [18].

To rank the interventions, we reported the surface under the cumulative ranking curve (SUCRA) scores. This score reflects the likelihood that an intervention is better than the competing interventions. To visualise the uncertainty in the network analyses, we also produced network forest plots. Heterogeneity and inconsistency were evaluated using τ^2, I^2 statistics and Cochran's Q test. We evaluated the consistency between direct and indirect evidence through both local and global approaches. Analysis of heterogeneity and node-splitting methods along with Q statistics to assess homogeneity and consistency were applied for this purpose. We performed assessment of selective publication by small sample bias methods for those outcomes with 10 or more studies. Funnel plots were constructed, and Egger's tests performed, to check for plot asymmetry. The threshold of significance was set at $p < 0.1$ for this method as this test has low power. Where asymmetries were present, a Duval and Tweedie's trim-and-fill procedure was applied to estimate bias-corrected effects.

Sensitivity analyses were undertaken by both sub-group and meta-regression analyses. Sub-group assessments were performed using either mixed-effects or random-effects models, where appropriate, for outcomes with 10 or more studies available. Meta-regressions of single features, one at a time, as well as multiple meta-regression with maximum likelihood estimator for τ^2 were conducted only for outcomes with 10 or more studies available per covariate. The multiple meta-regression models were submitted to a permutation test to confirm statistical significance. Both sub-group and meta-regression analyses were performed with a priori hypotheses – attempting to avert spurious associations – with the following features: operator experience; intubation technique; patient characteristics; setting; nature of procedure; type of laryngoscope; induction agent; opioid; and neuromuscular blocking drug used. Risk of bias judgements were planned but not performed throughout the sensitivity analyses since we did not have any study at low risk of bias. Additionally, we applied trial sequential analysis for the main outcome to deal with the risk of type-1 and type-2 errors due to repeated significance testing by subsequent meta-analyses.

Network meta-regressions were conducted, accounting for setting and predicted difficult airway. To assess the quality of the evidence for all outcomes, we applied the confidence in network meta-analysis (CINeMA), based on the grading of recommendations, assessment, development and evaluation (GRADE) approach [19–22]. This approach accounted for within-study risk of bias, heterogeneity between studies, indirectness of evidence, imprecision, publication bias and incoherence between direct and indirect evidence.

© 2021 Association of Anaesthetists

Results

Our initial search identified 50,243 articles with 33,511 remaining, after deduplication, for screening on title and abstract. A further 32,992 articles were excluded during such screening steps. The observed inter-rater agreement during the title and abstract screening was 98.5% (Cohen's $\kappa = 0.58$; 95% CI: 0.54–0.61; $p < 0.0001$). We could not retrieve 40 articles for full-text assessment. We excluded 298 articles during the full-text evaluation with 181 remaining included in the list of our systematic reviews. We did not analyse two trials which made data available only for crossover groups and hence we have reported the results of 179 studies. The observed inter-rater agreement during the full-text screening was 90.4% (Cohen's $\kappa = 0.79$; 95% CI: 0.74–0.85; $p < 0.0001$). Reasons why studies were excluded or not included are presented in Fig. 1. The full list of the included studies is presented in online Supporting Information (Appendix S1).

We included only randomised clinical trials which originated from 35 different countries. The studies involved patients from the general surgical population (n = 123); pregnant women (n = 3); elderly patients (n = 2); obese (n = 16); and patients with cervical immobilisation (n = 35). We judged whether studies' targeted populations were at particular risk of difficult intubation; most studies included patients not predicted to be at high risk of difficult intubation (n = 97), while some included specific conditions that might lead to difficulty, such as upper airway tumours and cervical immobilisation (n = 69). Some studies (n = 13) did not present sufficient data to be classified. Most studies evaluated patients undergoing elective procedures (n = 163), some assessed patients under urgent conditions (n = 11), and some did not specify (n = 5). Most of the studies were conducted in the operating theatre (n = 167), with a few in the Emergency Department (n = 3), ICU (n = 4), out of hospital (n = 4) and at multiple settings (n = 1). Most studies evaluated patients undergoing a standard intubation technique, that is induction, then facemask ventilation, then intubation (n = 154), while eight assessed patients intubated under rapid sequence induction; one included patients with both techniques, and 16 did not mention the technique used. Diverse induction techniques were applied throughout the studies with different combinations of hypnotics (propofol, thiopental, ketamine, etomidate, midazolam), opioids (fentanyl, sufentanil, remifentanil, morphine) and neuromuscular

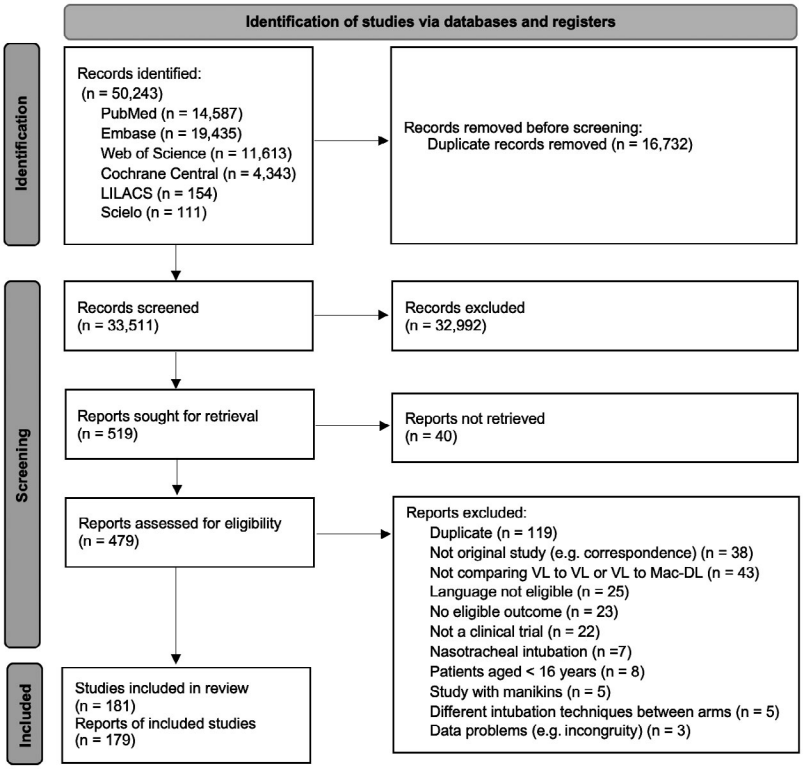

Figure 1 Review flow diagram. We did not analyse two included clinical trials with data available only from crossover arms and hence we reported data on 179 studies. VL, videolaryngoscope; Mac-DL, Macintosh direct laryngoscope.

© 2021 Association of Anaesthetists

blocking drugs (rocuronium, vecuronium, atracurium, cisatracurium, succinylcholine), as well as different doses and infusion techniques (bolus vs. target-controlled infusion).

We assessed the risk of bias at outcome level. A total of 808 outcome judgements were performed, with 513 (63.5%) being regarded at 'high risk', 295 (36.5%) at 'some concerns' and none at 'low risk' of overall bias according to Risk of Bias 2 tool [16]. The major concerns were related to measurement of outcomes, since it was not possible to blind operators or assessors to the devices used (Fig. 2).

The results of the individual studies for all outcomes are presented in the online Supporting Information (Figures S1–S3) and the published analytical code (https://rpubs.com/clistenescarvalho/Anaesthesia). In total, we assessed 20 different devices throughout the analyses: Airtraq™; Airtraq™ non-channelled; both AP Advances™ with Macintosh blade (APA MAC) and difficult airway blade (APA DAB); the three C-MAC® with Macintosh (C-MAC), Miller (C-MAC Miller) and D-Blade™ (C-MAC D); CEL-100™; Glidescope®; Imago V-blade®; both KingVisions® - channelled and non-channelled; Macintosh laryngoscope; McGrath™, both with Macintosh blade (McGrath MAC) and series 5; Pentax AWS™ both with Macintosh (Pentax AWS) and Miller blades (Pentax AWS Miller); Tosight™; Truview™; and UESCOPE® with Macintosh blade (UESCOPE MAC).

For any of our outcomes, we did not include in our network analyses those videolaryngoscopes with a single comparison available and no event. More details for such comparisons are presented in the online Supporting Information (Appendix S1). The Storz V-MAC® and the Storz DCI® were pooled together with the C-MAC for all outcomes. The McGrath series 3™ was included in the McGrath MAC node. A summary of the network geometry is shown in Figure 3. The SUCRA values and ranking of the devices for all outcomes are presented in Figure 4, while the league tables of the results for all pairs of comparisons are presented in the online Supporting Information (Tables S1–S7).

Failed intubation

The criteria for declared failed intubation varied among the studies, with the most frequent definition being more than two fails with the device. There were 120 studies comparing a videolaryngoscope with the Macintosh direct laryngoscope. We did not include 64 studies in the pairwise meta-analysis as there were no failed intubations in either group. Altogether, videolaryngoscopes were at lower risk of failed intubation than the Macintosh direct laryngoscope, with a risk ratio (95%CI) of 0.41 (0.29–0.58); $p < 0.0001$; 56 studies; 6396 participants; $I^2 = 17.5\%$, Chi^2-p = 0.1349. Trial sequential analysis confirmed the superiority of videolaryngoscopes over the Macintosh direct laryngoscope for this outcome (see online Supporting Information, Figure S4). More details about the pairwise analyses of this and other outcomes are presented in the online Supporting Information (Appendix S1). The pooled probability (95%CI) of failed intubation with the Macintosh direct laryngoscope was 0.86 (0.46–1.62)%. This means the successful intubation rate of the Macintosh direct laryngoscope is around 99.14%, whereas for videolaryngoscopes, it is around 99.65%.

We included 141 studies in the network meta-analysis, evaluating 16,478 patients for 16 different devices. Three videolaryngoscopes (C-MAC, Airtraq and Glidescope) presented a statistically significant lower risk of failed intubation compared with the Macintosh direct laryngoscope (Fig. 5).

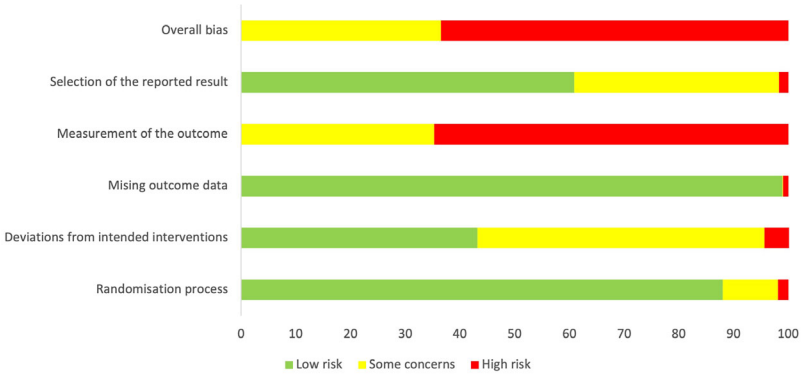

Figure 2 Risk of bias for multiple outcomes according to Risk of Bias 2 tool: summary of review authors' judgement about each domain, presented as percentages across included studies.

© 2021 Association of Anaesthetists

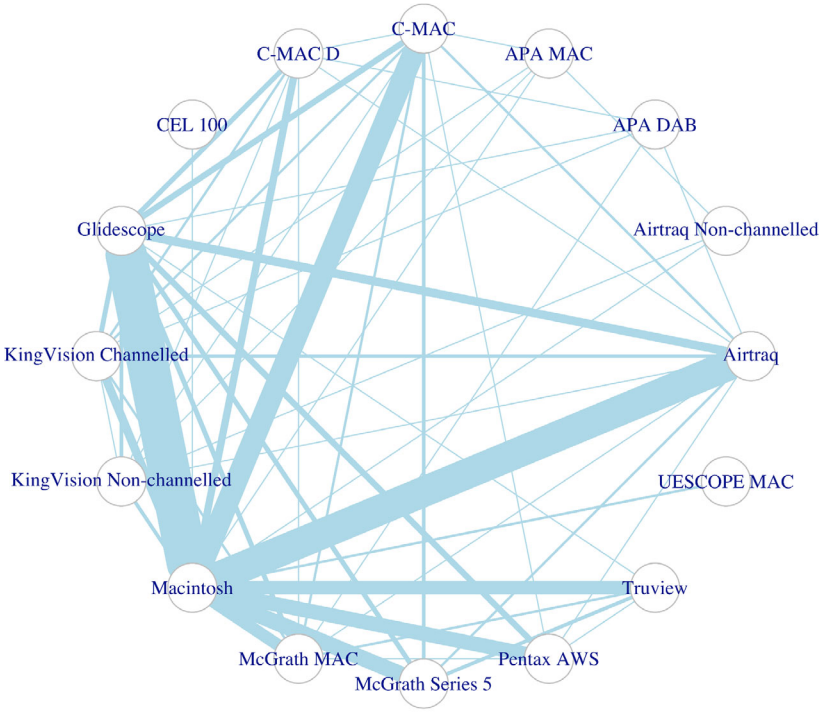

Figure 3 Network graph from the network meta-analysis comparing videolaryngoscopes and the Macintosh direct laryngoscope for failed intubation. Each node represents what we regarded as an individual intervention. Lines represent the direct comparisons available between the different devices and their thicknesses are proportional to the number of included clinical trials for the relevant comparisons. Abbreviations for videolaryngoscopes are defined in the text.

Failed first intubation attempt

There were 110 studies comparing videolaryngoscopes with the Macintosh direct laryngoscope. We did not include 12 studies in the pairwise meta-analysis as there were no events in either group. Failed first intubation attempt was less likely with the videolaryngoscopes than the Macintosh direct laryngoscope, with a risk ratio (95%CI) of 0.58 (0.47–0.71); $p < 0.0001$; 98 studies; 11,287 participants; $I^2 = 61.5\%$, Chi^2-$p < 0.0001$. The pooled probability (95% CI) of failed first intubation attempt with the Macintosh direct laryngoscope was 10.42 (8.17–13.2)%. Thus, the rate of successful first attempt intubation with the Macintosh direct laryngoscope is around 89.6%, whereas for videolaryngoscopes, it is around 94%.

We included 143 studies in the network meta-analysis, evaluating 16,953 patients for 19 devices. Four videolaryngoscopes (C-MAC, McGrath MAC, Airtraq and Glidescope) were statistically significantly less likely to lead to failed first intubation attempt compared with the Macintosh direct laryngoscope, whereas one (APA DAB) performed significantly worse (Fig. 5). The result for the APA DAB, however, relies on a single study [23].

Failed intubation within two attempts

This outcome differs from the failed intubation since most patients could still be intubated with the device in more than two attempts. Comparisons between videolaryngoscopes and the Macintosh direct laryngoscope were undertaken in 93 studies. We did not include 48 studies in the pairwise meta-analysis because they presented zero events in both groups. Videolaryngoscopes were at lower risk of failed intubation within two attempts when compared with the Macintosh direct laryngoscope, with a risk ratio (95% CI) of 0.53 (0.35–0.78); $p = 0.0022$; 45 studies; 5369 participants; $I^2 = 27\%$, Chi^2-$p = 0.0517$. The pooled probability (95% CI) of failing to intubate patients within two attempts with the Macintosh direct laryngoscope was 1.09 (0.59–1.99)%. This means the rate of successful intubation within two attempts with the Macintosh direct laryngoscope is around 99%, whereas for videolaryngoscopes, it is around 99.4%.

We included 120 studies in this network meta-analysis, evaluating 12,480 patients for 15 devices. Two videolaryngoscopes (C-MAC and APA MAC) presented a statistically significant lower risk of failed intubation within two attempts than the Macintosh direct laryngoscope (Fig. 5).

© 2021 Association of Anaesthetists

	Failed intubation	Failed first attempt	Failed 2 attempts	Difficult intubation	POGO	Difficult laryngoscopy	Time for intubation
C-MAC	78%	79%	74%	62%	57%	29%	75%
C-MAC D	78%	59%	38%	68%	43%	93%	66%
KingVision non-channelled	70%	69%	78%	63%	62%	82%	54%
McGrath MAC	64%	62%	33%	41%	44%	48%	37%
Pentax AWS	62%	49%	62%	87%	NA	77%	63%
Airtraq non-channelled	61%	73%	NA	NA	NA	71%	NA
Airtraq	60%	76%	52%	85%	74%	68%	81%
Truview	60%	50%	36%	20%	50%	80%	15%
Glidescope	59%	54%	38%	48%	36%	64%	36%
APA MAC	50%	48%	99%	NA	NA	20%	NA
CEL-100	41%	66%	68%	NA	NA	49%	NA
UESCOPE MAC	38%	49%	NA	NA	NA	19%	79%
McGrath series 5	33%	48%	14%	30%	51%	53%	5%
Macintosh	23%	24%	24%	8%	7%	11%	51%
KingVision channelled	16%	28%	43%	67%	76%	46%	59%
APA DAB	8%	3%	NA	NA	NA	15%	NA
Tosight	NA	64%	NA	NA	NA	NA	61%
Imago V-Blade	NA	27%	56%	21%	NA	57%	NA
Pentax AWS Miller	NA	24%	37%	NA	NA	41%	16%
C-MAC Miller	NA	NA	NA	NA	NA	27%	NA

Figure 4 'Heat map' of surface under the cumulative ranking curve (SUCRA) values of every device for all outcomes. The values may range from 0% to 100% and the higher the percentage, the higher the probability that the device ranks first or is in one of the top ranks. Highest SUCRA values are green, lowest ones are red. Interventions are sorted according to decreasing SUCRA values for the main outcome (failed intubation). POGO, percentage of glottic opening seen. Abbreviations for videolaryngoscopes are defined in the text.

© 2021 Association of Anaesthetists

Difficulty of intubation

Different scales were used to classify difficulty of intubation – most frequently the intubation difficulty scale [24]. Comparisons between videolaryngoscopes and the Macintosh direct laryngoscope were undertaken in 37 studies. One study showed zero events in both groups and therefore was not included in the pairwise meta-analysis. Difficult intubation was less likely that with the Macintosh direct laryngoscope, with a risk ratio (95%CI) of 0.48 (0.38–0.61); $p < 0.0001$; 36 studies; 3037 participants; $I^2 = 84.1\%$, Chi^2-$p < 0.0001$. The pooled probability (95% CI) of even a slightly difficult intubation with the Macintosh direct laryngoscope was 54.88 (35.62–72.79)%. Hence, the rate of easy intubation with the Macintosh direct laryngoscope is around 45.1%, whereas for videolaryngoscopes, it is around 73.7%.

We included 49 studies in this network meta-analysis, evaluating 4409 patients for 12 devices. Seven videolaryngoscopes (C-MAC, C-MAC D, both KingVisions channelled and non-channelled, Pentax AWS, Airtraq and Glidescope) presented a statistically significant lower risk of even a slightly difficult intubation (intubation difficulty scale > 0) as compared with the Macintosh direct laryngoscope (Fig. 5).

Improved glottic view

This was expressed as the percentage of glottic opening seen. Comparisons between videolaryngoscopes and the Macintosh direct laryngoscope were conducted in 25 studies, with 31 comparisons. We did not include 11 comparisons in the meta-analyses either because they presented either data on medians or we could not extract appropriate data. In four of these 11 comparisons, there was no evidence of a significant difference between the videolaryngoscopes and the Macintosh direct laryngoscope; seven suggested benefit from the use of the videolaryngoscopes. Pooled results demonstrated an improved glottic view for videolaryngoscopes compared with the Macintosh direct laryngoscope, with a standardised mean difference (95%CI) of 1.21 (0.82–1.60); $p < 0.0001$; 20 comparisons; 1870 participants; $I^2 = 92.8\%$, Chi^2-$p < 0.0001$. The pooled mean percentage of glottic opening (95%CI) for the Macintosh direct laryngoscope was 65.50 (56.94–74.10)%, whereas for videolaryngoscopes, it was 86.77 (82.42–91.11)%.

For the network assessment, we included 31 studies evaluating 5135 patients for 15 different devices. We had 43 comparisons between different devices, but 19 comparisons (10 studies) were not included in the network analysis because they either presented data on

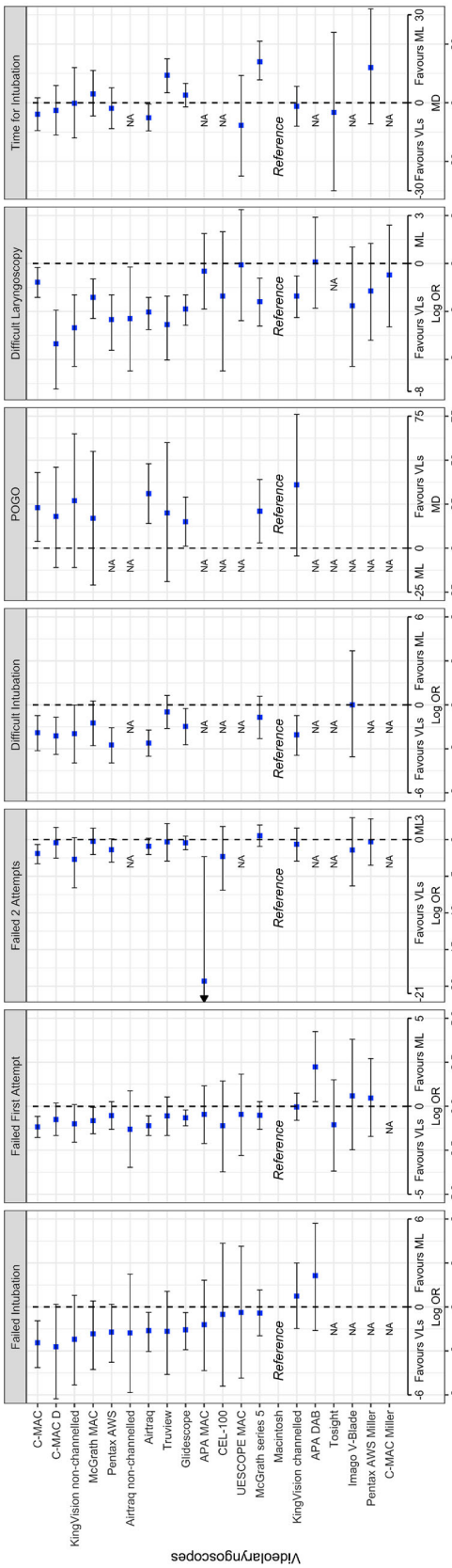

Figure 5 Forest plots for the seven assessed outcomes with the comparative network effect sizes of all interventions: data are sorted by descending SUCRA values for the main outcome (failed intubation). Blue squares represent the estimated network effect sizes. Black bars represent the 95% credible intervals (95% CrIs). Values are number for time to intubation (s) and proportion (percentage for percentage of glottic opening, POGO). ML, Macintosh laryngoscope; VL, videolaryngocsope; POGO, percentage of glottic opening; NA, not available; OR, odds ratio; MD, mean difference. Abbreviations for videolaryngoscopes are defined in the text.

© 2021 Association of Anaesthetists

medians or we could not extract appropriate data. Details about these studies are presented in the online Supporting Information (Appendix S1). The 21 remaining studies evaluated 2115 patients for 10 devices. Four videolaryngoscopes (C-MAC, Airtraq, Glidescope and McGrath series 5) had a statistically significant higher percentage of glottic opening scores compared with the Macintosh direct laryngoscope (Fig. 5).

Difficult laryngoscopy by Cormack and Lehane classification

We extracted data comparing videolaryngoscopes with the Macintosh direct laryngoscope from 92 studies. We did not include 12 studies in the pairwise meta-analysis because they presented zero events in both groups. Videolaryngoscopes significantly reduced the risk of difficult laryngoscopy compared with the Macintosh direct laryngoscope, with a risk ratio (95% CI) of 0.24 (0.19–0.31); $p < 0.0001$; 80 studies; 9324 participants; $I^2 = 51\%$, Chi^2-$p < 0.0001$. The pooled probability (95%CI) of difficult laryngoscopy with the Macintosh direct laryngoscope was 12.65 (9.73–16.30)%. This implies a rate of easy laryngoscopy with the Macintosh direct laryngoscope of around 87.4%, and around 97% for videolaryngoscopes.

We included 117 studies in this network meta-analysis, evaluating 13,749 patients for 19 devices. Eleven videolaryngoscopes presented a statistically significant lower risk of difficult laryngoscopy in comparison with the Macintosh direct laryngoscope (Fig. 5).

Time for intubation

Definitions of this outcome varied and included the time at which the tracheal tube was seen to pass through the vocal cords, and the appearance of exhaled carbon dioxide on capnography. Comparisons between videolaryngoscopes and the Macintosh direct laryngoscope were conducted in 134 studies with 169 comparisons. We did not include 43 comparisons in the meta-analyses either because they presented data on medians or did not present information on data dispersion. Out of these 43 comparisons, 21 did not find evidence of a significant difference between the videolaryngoscopes and the Macintosh direct laryngoscope; the Macintosh direct laryngoscope was faster in 19 and the videolaryngoscopes were better in three. After summarising the results of those studies reporting means with measures of dispersion, we found videolaryngoscopes took longer than the Macintosh direct laryngoscope for intubation, with a standardised mean difference (95% CI) of 0.29 (0.10–0.47); $p = 0.0023$; 126 comparisons; 10,613 participants; $I^2 = 94.7\%$, Chi^2-$p < 0.0001$. The pooled mean time (95% CI) for intubation with the Macintosh direct laryngoscope was 33.06 (29.32–36.79) s, whereas for videolaryngoscopes, it was 34.87 (31.40–38.34) s.

For the network assessment, we included 167 studies evaluating 21,005 patients for 24 different devices. We had 213 comparisons between different devices, but 63 (45 studies) were not included either because they presented data as medians or we could not extract relevant data. Details about these studies are presented in the online Supporting Information (Appendix S1). The 122 remaining studies evaluated 12,268 patients for 14 devices. We found the Airtraq to be significantly quicker, and the McGrath series 5 and the Truview significantly slower, than the Macintosh laryngoscope, but the 95% credible intervals for the comparisons suggested that these differences are not clinically relevant (Fig. 5).

We did not find evidence of significant inconsistency between direct and indirect evidence based on the global approach for our main outcome (failed intubation), but we did for some comparisons. As the inconsistent comparisons were few and had no relevant influence on the comparative effect sizes and the ranking of the interventions, we did not downgrade the overall quality of the evidence for our main outcome for incoherence. Significant inconsistency for both the global and local approaches was found for failed first intubation attempt and difficulty of intubation. Only a few comparisons presented inconsistency between direct and indirect evidence for the remaining outcomes.

The risks of publication bias were investigated by standard pairwise meta-analyses including all pairs of devices. We found possible selective reporting for: failed first intubation attempt; difficult intubation; percentage of glottic opening; difficult laryngoscopy; and time for intubation. Sub-group and sensitivity analyses for the pairwise meta-analyses are presented in the online Supporting Information (Appendix S1).

A sensitivity analysis for setting showed some influence of this feature over the heterogeneity between direct and indirect evidence for failed intubation. However, the setting did not materially influence the comparative effect sizes and the ranking of the interventions for this outcome. Rankings from the sub-group analyses by operator experience and predicted difficult intubation are presented in Figures 6 and 7, respectively.

The overall quality of the evidence was regarded as low for failed intubation, and for failed intubation within two attempts. It was considered very low quality for the remaining outcomes: failed first intubation attempt; difficulty of intubation; percentage of glottic opening; difficult laryngoscopy; and time for intubation. The judgements of

© 2021 Association of Anaesthetists

CINeMA domains for each comparison for failed intubation are presented in online Supporting Information (Table S8).

Discussion

We have found a statistically and clinically significant reduction in failed intubation with three videolaryngoscopes compared with the Macintosh direct laryngoscope in a diverse population of adult patients undergoing orotracheal intubation for both elective and urgent procedures, in different settings. These were the C-MAC, the Airtraq and the Glidescope. Of these, the C-MAC was ranked highest for failed intubation and was the only device to be statistically significantly better in comparison with the Macintosh direct laryngoscope for all outcomes. The overall quality of the evidence for this finding was moderate (see online Supporting Information, Table S8). However, the C-MAC and the C-MAC D-Blade were the devices presenting the highest probabilities of being the best laryngoscopes for successful intubation (Fig. 4) and might therefore figure among the preferred videolaryngoscopes.

We also found that the C-MAC and C-MAC D-Blade were most likely to be the best videolaryngoscopes for intubation success when used by experienced operators, and that the Macintosh videolaryngoscopes were best when used by non-experienced operators (Fig. 6). This is to be expected, since most 'airway managers' are more used to Macintosh blades. On the other hand, the C-MAC D-Blade was much more likely to be better than the C-MAC for failed intubation in patients predicted to have difficult intubation. We might therefore benefit from the availability of hyperangulated videolaryngoscopes such as the C-MAC D-Blade, bearing in mind that continuous training is necessary to acquire and maintain expertise with such devices.

However, we could not identify statistically significant differences between the various videolaryngoscopes assessed for failed intubation in any of the evaluated scenarios when accounting for the uncertainty across the analyses (Fig. 5), despite the rankings obtained. Also, the pairwise analyses pooling together all the videolaryngoscopes and comparing them with the Macintosh direct laryngoscope for failed intubation did not present significant heterogeneous effect sizes, suggesting there is similarity between the videolaryngoscopes. Thus, we are unable to suggest that any single videolaryngoscope, or small group of scopes, may be better for preventing failed intubation, nor can we be sure that any scope is not advised for this purpose. Additionally, most videolaryngoscopes were ranked above the Macintosh direct laryngoscope for failed intubation, and the overall superiority of videolaryngoscopes

	Operator experience	
Laryngoscopes	Experienced	Non-experienced
C-MAC D	75%	NA
C-MAC	74%	82%
KingVision non-channelled	69%	50%
Pentax AWS	67%	2%
Glidescope	67%	50%
McGrath MAC	62%	77%
Airtraq	59%	54%
Airtraq non-channelled	55%	NA
McGrath series 5	52%	18%
CEL-100	39%	NA
UESCOPE MAC	37%	NA
Truview	37%	76%
APA MAC	37%	72%
KingVision channelled	28%	18%
Macintosh	24%	52%
APA DAB	17%	NA

Figure 6 'Heat map' of surface under the cumulative ranking curve (SUCRA) values of different devices for failed intubation by sub-groups of experienced and non-experienced operators. The higher the percentage, the higher the probability that the device ranks first or is in one of the top ranks. Highest SUCRA values are green, lowest are red. Interventions are sorted according to decreasing SUCRA values for the main outcome (failed intubation). Abbreviations for videolaryngoscopes are defined in the text.

	Predicted difficult intubation	
Laryngoscopes	Predicted	Non-predicted
Truview	83%	59%
McGrath MAC	77%	NA
C-MAC D	76%	78%
Airtraq	72%	27%
KingVision non-channelled	60%	69%
Glidescope	58%	43%
KingVision channelled	47%	6%
McGrath series 5	45%	30%
Airtraq non-channelled	43%	NA
Pentax AWS	37%	95%
C-MAC	37%	73%
APA MAC	28%	54%
APA DAB	22%	NA
Macintosh	14%	29%
McGrath MAC	NA	47%
CEL 100	NA	39%

Figure 7 'Heat map' of surface under the cumulative ranking curve (SUCRA) values of different devices for failed intubation for predicted and non-predicted difficult intubation. The higher the percentage, the higher the probability that the device ranks first or is in one of the top ranks. Highest SUCRA values are green, lowest are red. Interventions are sorted according to decreasing SUCRA values for the main outcome (failed intubation). Abbreviations for videolaryngoscopes are defined in the text.

© 2021 Association of Anaesthetists

was confirmed by convincing results from the pairwise analyses. We feel that the statistically significant results reached by the C-MAC, the Airtraq and the Glidescope for failed intubation may simply reflect the greater body of data available for these devices, and it is likely that other videolaryngoscopes are also less likely to lead to failed intubation than the Macintosh direct laryngoscope. Consequently, there is still considerable uncertainty about the order displayed, and further research comparing different videolaryngoscopes would be worthwhile.

We have also found reliable evidence that videolaryngoscopes improve glottic view in comparison with the Macintosh direct laryngoscope, both by significantly reducing the risk of difficult laryngoscopy (Cormack and Lehane grades 3 or 4) and enhancing the percentage of glottic opening scores. Another important finding from the network analyses was the significantly lower chance of facing a difficult laryngoscopy with the C-MAC D-Blade compared with the C-MAC Macintosh blade. This demonstrates the improvement in glottic view that is possible with a hyperangulated blade when the view is poor with Macintosh videolaryngoscopes. We should bear in mind, however, that having a better view does not necessarily lead to successful intubation, as suggested by the lack of evidence for a difference in intubation-related outcomes between hyperangulated and less angulated blades. Accordingly, videolaryngoscopes increase the chances of intubating the trachea with fewer attempts in comparison with the Macintosh direct laryngoscope, with lower risk of both failed first intubation attempt and failed intubation within two attempts. However, we should consider the large heterogeneity from these analyses and the significantly worse performance found for the hyperangulated APA DAB for failed first attempt. Along with the results from the other outcomes, it may highlight the differential performance of different devices under different conditions, although we should be cautious interpreting the APA DAB result since it relies on a single study [23]. Similarly, videolaryngoscopes reduced the risk of difficult intubation compared with the Macintosh direct laryngoscope, with seven videolaryngoscopes performing significantly better than the Macintosh direct laryngoscope for this outcome.

Apart from the improvements brought by videolaryngoscopes to airway management, concerns have been raised about the time taken to accomplish tracheal intubation with these devices [10, 11]. Delays might lead to deterioration in a patient's clinical condition, especially in urgent procedures such as rapid sequence induction. However, we did not find clinically relevant differences between any videolaryngoscope and the Macintosh direct laryngoscope for time needed to intubate patients according to the estimated 95% credible intervals, even though little statistically significant delay has been found for videolaryngoscopes. Furthermore, all 126 mean differences available for time to intubation between videolaryngoscopes and the Macintosh direct laryngoscope were within the range of −40 to 70 s. We can, therefore, by counting these results along with those from the other intubation-related outcomes, infer clinical equivalence between all videolaryngoscopes and the Macintosh direct laryngoscope for time to intubation, despite the heterogeneous results and the different definitions used for this outcome.

Our results for time to intubation are in accordance with those of the only network analysis we have found comparing videolaryngoscopes, and other devices, with the Macintosh direct laryngoscope [9]. This review evaluated only patients with neck immobilisation and, likewise, demonstrated the 95% credible intervals of mean time difference between different videolaryngoscopes and the Macintosh direct laryngoscope to be within a clinically equivalent range. The authors also demonstrated that some videolaryngoscopes outperformed the Macintosh direct laryngoscope for first-pass success and glottic view, but could not find any significant difference between any two different videolaryngoscopes for such outcomes.

Many other systematic reviews with pairwise meta-analyses comparing videolaryngoscopes with the Macintosh direct laryngoscope are available and demonstrate videolaryngoscopes' superiority in relation to the Macintosh direct laryngoscope for different outcomes across different populations [8, 11–14, 25]. Some of them also point to differential performance between different videolaryngoscopes [8, 14]. However, these reviews did not compare the videolaryngoscopes individually and could also not arrange rankings for the devices.

Another relevant inference drawn from our results is the role of parallel covariates in the videolaryngoscopes' performance – as demonstrated throughout the sensitivity and sub-group analyses (see online Supporting Information, Appendix S1). Characteristics such as operator experience significantly improved videolaryngoscopes' performance, highlighting the importance of training to make the best use of equipment. The relative risk of failed intubation with videolaryngoscopes in comparison with the Macintosh direct laryngoscope for patients predicted to have difficult intubation was half that of patients not predicted to be difficult, although the difference did not attain statistical significance. It is likely, therefore, that different devices match better to specific scenarios and thus complement one another [26]. The videolaryngoscopes

© 2021 Association of Anaesthetists

with hyperangulated blades, for example, may be more difficult to manipulate, need more training and may perhaps hamper airway management in those patients with easy airways. Conversely, they may improve glottic view in patients with an impaired glottic view with Macintosh videolaryngoscopes and even enable tracheal intubation in such patients. We might then benefit from the clinical availability of videolaryngoscopes with both types of blades, and acquiring and maintaining skills with both.

Our results have some limitations. The overall quality of the evidence supporting the rankings presented was low or very low. Additionally, we had too few comparisons available for some devices such as APA DAB, Imago V-blade, C-MAC Miller, Pentax AWS Miller, Tosight, UESCOPE MAC and Airtraq non-channelled. Caution should therefore be taken when interpreting their results. We also did not assess the risk of minor and major complications; clearly an important point to be taken into account. The way we clustered the interventions (the network nodes) may also have had some influence over our results – possibly impairing the performance of some videolaryngoscopes as well as preventing the reach of statistical significance in some cases. Additionally, there is emerging evidence that a proportion of published randomised trials present data problems such as false data [27]. This may have affected our results, in common with any systematic review and meta-analysis, but it is difficult to speculate how.

In conclusion, we have found convincing evidence that videolaryngoscopes outperform the Macintosh direct laryngoscope for many outcomes in orotracheal intubation in a range of adult patients. Some devices were more evenly top ranked throughout the evaluated outcomes and scenarios and may be currently preferred among the available videolaryngoscopes. Furthermore, videolaryngoscopes and the Macintosh direct laryngoscope are also clinically equivalent for the time taken to accomplish tracheal intubation. However, despite differential performances found between the devices, we could not identify statistically significant superiority of one videolaryngoscope, or a small set of them, over the other ones available and, therefore, further research comparing different types of videolaryngoscopes might help refine the rankings we have presented.

Acknowledgements

The review protocol was registered on PROSPERO (CRD42021220345) and published separately (Carvalho et al, preprint, https://doi.org/10.1101/2021.01.18.21250062). No external funding or competing interests declared.

References

1. 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.
2. 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.
3. Roth D, Pace NL, Lee A, et al. Bedside tests for predicting difficult airways: an abridged Cochrane diagnostic test accuracy systematic review. *Anaesthesia* 2019; **74**: 915–28.
4. Detsky ME, Jivraj N, Adhikari NK, et al. Will this patient be difficult to intubate? The rational clinical examination systematic review. *Journal of the American Medical Association* 2019; **321**: 493–503.
5. Nørskov AK, Rosenstock CV, Wetterslev J, Astrup G, Afshari A, Lundstrøm LM. 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.
6. Gibbins M, Kelly FE, Cook TM. Airway management equipment and practice: time to optimise institutional, team, and personal preparedness. *British Journal of Anaesthesia* 2020; **125**: 221–4.
7. Brown CA, Kaji AH, Fantegrossi A, et al. Video versus augmented direct laryngoscopy in adult emergency department intubations. *British Journal of Anaesthesia* 2020; **125**: e199–200.
8. Lewis SR, Butler AR, Parker J, Cook TM, Schofield-Robinson OJ, Smith AF. Videolaryngoscopy versus direct laryngoscopy for adult patients requiring tracheal intubation: a Cochrane Systematic Review. *British Journal of Anaesthesia* 2017; **119**: 369–83.
9. Singleton BN, Morris FK, Yet B, Buggy DJ, Perkins AB. Effectiveness of intubation devices in patients with cervical spine immobilisation: a systematic review and network meta-analysis. *British Journal of Anaesthesia* 2021; **126**: P1055–66.
10. Kaplan MB, Hagberg CA, Ward DS, et al. Comparison of direct and video-assisted views of the larynx during routine intubation. *Journal of Clinical Anesthesia* 2006; **18**: 357–62.
11. Hoshijima H, Mihara T, Maruyama K, et al. C-MAC videolaryngoscope versus Macintosh laryngoscope for tracheal intubation: a systematic review and meta-analysis with trial sequential analysis. *Journal of Clinical Anesthesia* 2018; **49**: 53–62.
12. Hoshijima H, Mihara T, Maruyama K, et al. McGrath videolaryngoscope versus Macintosh laryngoscope for tracheal intubation: a systematic review and meta-analysis with trial sequential analysis. *Journal of Clinical Anesthesia* 2018; **46**: 25–32.
13. Hoshijima H, Denawa Y, Tominaga A, Nakamura C, Shiga T, Nagasaka H. Videolaryngoscope versus Macintosh laryngoscope for tracheal intubation in adults with obesity: a systematic review and meta-analysis. *Journal of Clinical Anesthesia* 2018; **44**: 69–75.
14. Pieters BMA, Maas EHA, Knape JTA, van Zundert AAJ. Videolaryngoscopy vs. direct laryngoscopy use by experienced anaesthetists in patients with known difficult airways: a systematic review and meta-analysis. *Anaesthesia* 2017; **72**: 1532–41.
15. Thomas J, Brunton J, Graziosi S. *EPPI-Reviewer 4.11.5.1: software for research synthesis. EPPI-Centre Software*. London, UK: Social Science Research Unit, Institute of Education, University of London; 2010.

 © 2021 Association of Anaesthetists

16. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *British Medical Journal* 2019; **366**: l4898.
17. Van Valkenhoef G, Lu G, de Brock B, Hillege H, Ades AE, Welton NJ. Automating network meta-analysis. *Research Synthesis Methods* 2012; **3**: 285–99.
18. Chaimani A, Caldwell DM, Li T, et al. Undertaking network meta-analyses. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.2*. The Cochrane Collaboration 2021. www.training.cochrane.org/handbook (accessed 01/08/2021).
19. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, et al. Cinema: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Medicine* 2020; **17**: e1003082.
20. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal* 2008; **336**: 924–6.
21. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is "quality of evidence" and why is it important to clinicians? *British Medical Journal* 2008; **336**: 995–8.
22. Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. *British Medical Journal* 2008; **336**: 1049–51.
23. Kleine-Brueggeney M, Greif R, Schoettker P, Savoldelli GL, Nabecker S, Theiler LG. Evaluation of six videolaryngoscopes in 720 patients with a simulated difficult airway: a multicentre randomized controlled trial. *British Journal of Anaesthesia* 2016; **116**: 670–9.
24. Adnet F, Borron S, Racine S, et al. The Intubation Difficulty Scale (IDS) proposal and evaluation of a new score characterizing the complexity of endotracheal intubation. *Anesthesiology* 1997; **87**: 1290–7.
25. Liu TT, Li L, Wan L, Zhang CH, Yao WL. Videolaryngoscopy vs. Macintosh laryngoscopy for double-lumen tube intubation in thoracic surgery: a systematic review and meta-analysis. *Anaesthesia* 2018; **73**: 997–1007.
26. Norris A, Heidegger T. Limitations of videolaryngoscopy. *British Journal of Anaesthesia* 2016; **117**: 148–50.
27. Carlisle JB. False individual patient data and zombie randomised controlled trials submitted to *Anaesthesia*. *Anaesthesia* 2020; **76**: 472–9.

Supporting Information

Additional supporting information may be found online via the journal website.

Appendix S1. Further information for the pairwise and network analyses performed.

Figure S1. Forest plot for the comparison between videolaryngoscopes and Macintosh laryngoscope for failed intubation, sorted by increasing relative risk.

Figure S2. Forest plot for the comparison between videolaryngoscopes and Macintosh laryngoscope for failed first intubation attempt, sorted by increasing relative risk.

Figure S3. Forest plot for the comparison between videolaryngoscopes and Macintosh laryngoscope for time for intubation, sorted by increasing standardised mean difference.

Figure S4. Trial sequential analysis for the comparison between videolaryngoscopes and Macintosh laryngoscope for risk of failed intubation

Table S1. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for failed intubation.

Table S2. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for failed first intubation attempt.

Table S3. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for failed intubation within two attempts.

Table S4. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for difficult intubation.

Table S5. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for percentage of glottic opening.

Table S6. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for difficult laryngoscopy (Cormack and Lehane ≥3).

Table S7. League table with the estimated effect sizes for each comparison between two different devices included in the network meta-analysis of videolaryngoscopes and Macintosh laryngoscope for time to intubation.

Table S8. Report of the assessment for the Confidence in Network Meta-analysis (CINeMA) of videolaryngoscopes and Macintosh laryngoscope for failed intubation – a system based on the GRADE approach.

© 2021 Association of Anaesthetists

Next Generation SedLine® Brain Function Monitoring and O3® Regional Oximetry

Available Together on the Root® Platform

Root with Next Generation SedLine and O3 provides a more complete picture of the brain through an instantly interpretable and adaptable display.

- Next Generation SedLine helps clinicians monitor the state of the brain under anaesthesia with an enhanced signal processing engine and four leads of bilateral EEG
- O3 monitors cerebral oxygenation in situations where pulse oximetry alone may not be fully indicative of the oxygen in the brain

Discover Brain Function Monitoring

www.masimo.co.uk

© 2021 Masimo. All rights reserved.

For professional use. See instructions for use for full prescribing information, including indications, contraindications, warnings, and precautions.

Masimo

PLCO-0040048/PLLT-11085C-0620