PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	DirEct Versus VIdeo LaryngosCopE (DEVICE): Protocol and statistical analysis plan for a randomized clinical trial in critically ill adults undergoing emergency tracheal intubation
AUTHORS	Prekker, Matthew; Driver, Brian; Trent, Stacy; Resnick-Ault, Daniel; Seitz, Kevin; Russell, Derek; Gandotra, Sheetal; Gaillard, John; Gibbs, Kevin; Latimer, Andrew; Whitson, Micah; Ghamande, Shekhar; Vonderhaar, Derek; Walco, Jeremy; Hansen, Sydney; Douglas, Ivor S.; Barnes, Christopher; Krishnamoorthy, Vijay; Bastman, Jill; Lloyd, Bradley; Robison, Sarah; Palakshappa, Jessica; Mitchell, Steven; Page, David; White, Heath; Espinera, Alyssa; Hughes, Christopher; Joffe, AM; Herbert, J. Taylor; Schauer, Steven; Long, Brit; Imhoff, Brant; Wang, Li; Rhoads, Jillian; Womack, Kelsey; Janz, David; Self, Wesley; Rice, Todd; Ginde, Adit; Casey, Jonathan; Semler, Matthew

VERSION 1 – REVIEW

REVIEWER	Hansel , Jan The University of Manchester
REVIEW RETURNED	30-Oct-2022

GENERAL COMMENTS	It was an absolute pleasure reviewing the protocol for the DEVICE trial by Prekker et al., a proposed multi-centre randomized controlled trial comparing videolaryngoscopy to direct laryngoscopy in critically ill adults undergoing tracheal intubation. I have reviewed the clinicaltrials.gov registration (NCT05239195). Registered prospectively. Outcomes in CT.gov and proposed protocol matching. Methods and analysis: - The randomization process is clearly described and would be at low risk of bias. - The group should be congratulated on the adoption of a very pragmatic design by not mandating the brand of laryngoscope. I also commend the necessary distinction of hyperangulated and Macintosh-style/standard geometry VLs, but would just want to know how the various brands and blade types in common use will be allocated. For example a McGrath MAC blade is slightly more curved than a standard Macintosh blade, but would fit more comfortably into that category, of course. I also note no mention is made of channelled VLs. Are these in use in any of the participating sites? It might be worth explicitly mentioning that channelled VLs are not used, if that is indeed the case. - I note that in the VL arm operators are not required to use the video screen. This would in effect make an attempt where a video screen is not used a direct laryngoscopy attempt, and might

	muddy the waters a bit. An argument could be made that were a Macintosh-style VL used in such a way, it could be counted into the DL category. Will the frequencies of this type of VL use be reported on CRF and in the final manuscript? - The exclusion criterion of operator preference when VL/DL might be indicated/contraindicated is notable. Will the eCRF collect data such as reasons for exclusion for excluded participants, and this reported in the CONSORT flow diagram? - The primary outcome is sensible and explicitly defined as well. - The composite secondary outcome is interesting and captures an array of important patient-oriented outcomes. It would be helpful to report the four components separately as well as the composite incidences. - The definition of duration of laryngoscopy and tracheal intubation could be more robust. You state 'the final placement of an endotracheal tube or tracheostomy tube in the trachea'. Do you define the final placement as (a) insertion as declared by the intubator, (b) time to cuff inflation, (c) time to first capnography trace, or (d) other? It would be helpful to clarify this prior to data collection if possible, and if it is already defined, to have it more explicitly reported in the final manuscript. - The definition of oesophageal intubation is similarly contentious. Is it oesophageal intubation that is recognised visually or by virtue of an absent capnography trace or clinical signs? A clearer definition would be welcome as this is an important outcome associated with significant downstream morbidity when missed. - The definitions of ICU-FDs and VFDs are clear and explicit. - The remaining important outcomes are clearly defined. - I note the distinct absence of overall intubation failure as a reported outcome. It would be extremely valuable to have this collected, especially in this high-risk cohort, as it is arguably a more robust and patient-oriented outcome than first-pass success. We used the definition of more than three failed attempts as failed intubation in our Cochrane review. - The sample size calculation seems sensible. The absolute risk reduction of approximately 3% for failed intubation and approximately 5.5% for FPS (favouring VL) was noted in the recent Cochrane review. Provided that the remaining assumptions are realistic, it should result in a sufficiently powered study. - The stopping boundary is appropriately conservative. - I note the appropriate use of intention-to-treat analysis for the primary and secondary outcomes. Given the significant potential for crossover, will a per protocol analysis be conducted and reported as well? - I note the recommendations by the FDA and EMA for P value reporting, but do wonder whether it would make sense to report effect estimates with 95% CI for the primary and secondary outcome in the manuscript as well. This will aid future data extraction for evidence synthesis analyses. - The set of variables collected is comprehensive and clearly defined. - The secondary analyses are meaningfully defined and the predicted hypotheses in agreement with what is available in the literature to date. General comments and observations: - Consider replacing all instances of 'endotracheal' tube/intubation with 'tracheal' tube/intubation for uniformity with other published works on the topic. Orotracheal as a term, as used in the inclusion criteria, seems sensible as well.

	Major concerns: I have no major concerns about this protocol. This is an overall very robust protocol for what will surely be a well-conducted trial with important new knowledge generated in this area of airway management research. I would again like to congratulate the authors on this significant undertaking and wish them the best of luck in getting it done. I look forward to the final results.

REVIEWER	Loop, T Universitätsklinikum Freiburg, Anaesthesiology
REVIEW RETURNED	06-Dec-2022

GENERAL COMMENTS	In the manuscript, Prekker and co-authors present the study protocol for a prospective randomised trial of airway management. The hypothesis is that in the comparison of direct laryngoscopy vs. videolaryngoscopic assisted laryngoscopy differences in successful endotracheal intubation are to be achieved. Adverse events are particularly associated with emergency airway management and show the highest risk of complications. In this context, hypoxia is one of the most common causes of airway management-associated deaths. In this respect, the primary endpoint for this study is correctly chosen. The increasingly common use of videolaryngoscopes leads to new and additional risks for airway injuries, as rigid guide rods are often used. Therefore, airway injuries should be included as secondary endpoints in the study. The incidence of difficult mask ventilation is 1.5%, that of impossible mask ventilation 0.16%. The incidence of difficult direct laryngoscopy is 1.5% to 8.0% depending on the collective, the incidence of difficult intubation is slightly lower. An unexpected "cannot intubate, cannot ventilate" situation has a probability of 0.008% (1:13,000) to 0.004% (1:25,000). These incidences should also be recorded as they are significantly higher under emergency conditions. Hyperangulated spatulas are used. Can we assume that the video laryngoscope with hyperangulated blades are of the D-BLADE type? Or are different blades used? Since the study is multicentre, this is an important point. I doubt that, as the authors describe on page 20, it makes no difference whether intubation is done in the ED or in the ICU. The patient populations are very different, the risk of aspiration is higher, the oxygen reserve in the ICU is lower (as the indication is more likely to be given later), and so on.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Methods and analysis:

- The randomization process is clearly described and would be at low risk of bias.

Thank you for this comment.

- The group should be congratulated on the adoption of a very pragmatic design by not mandating the brand of laryngoscope. I also commend the necessary distinction of hyperangulated and Macintosh-style/standard geometry VLs, but would just want to know how the various brands and blade types in common use will be allocated. For example a McGrath MAC blade is slightly more curved than a standard Macintosh blade, but would fit more comfortably into that category, of course.

This is an astute observation. In the Study Interventions/Video Laryngoscope Group section (page 12) of the protocol, we intentionally keep the trial requirement rather general: “Trial protocol does not dictate the brand of video laryngoscope or the geometry of the laryngoscope blade (e.g., hyperangulated vs. non-hyperangulated), but these details will be recorded.”

In the study results manuscript, our plan is to report laryngoscope manufacturer and blade style in a table including characteristics of the intubation procedure, grouped as follows:

1. Direct laryngoscope
    a. Standard geometry blade shape
        i. Curved blade (Macintosh)
        ii. Straight blade (Miller)
2. Video laryngoscope
    a. Standard geometry (non-hyperangulated) blade shape
        i. Storz C-MAC
        ii. McGrath MAC
        iii. Glidescope MAC
    b. Hyperangulated blade shape
        i. GlideScope G/AVL
        ii. GlideScope LoPro
        iii. Storz D Blade

- I also note no mention is made of channelled VLs. Are these in use in any of the participating sites? It might be worth explicitly mentioning that channelled VLs are not used, if that is indeed the case.

While channeled video laryngoscopes (e.g., AirTraq, others) are not forbidden in the trial protocol, they are not in use at any of the trial sites, and we do not expect any trial participants to be intubated with one of these devices. We will provide a full list of all laryngoscopes used in the trial in the results manuscript.

- I note that in the VL arm operators are not required to use the video screen. This would in effect make an attempt where a video screen is not used a direct laryngoscopy attempt, and might muddy the waters a bit. An argument could be made that were a Macintosh-style VL used in such a way, it could be counted into the DL category. Will the frequencies of this type of VL use be reported on CRF and in the final manuscript?

Providers are encouraged, rather than required, to view the screen in the video laryngoscope group, as a safety precaution in case of screen failure or obscuration of the camera port by blood or emesis. We expect the video screen to be used in nearly all cases. Further, we capture data on whether the operator views the video screen during the intubation, which will be reported in the final manuscript, and which could (if a significant number of cases were to occur) be used in a post-hoc sensitivity analysis as suggested by the reviewer. That analysis has not been prespecified given our belief that too few such cases will occur to warrant this analysis.

- The exclusion criterion of operator preference when VL/DL might be indicated/contraindicated is notable. Will the eCRF collect data such as reasons for exclusion for excluded participants, and this reported in the CONSORT flow diagram?

Yes, we are capturing reasons for exclusion for the excluded patients and those data will be included in the CONSORT diagram with the manuscript. We agree with the reviewer that understanding our final trial population and interpretating of our trial results will require a careful accounting of the frequency and reasons for exclusion, especially when the reason is related to clinician lack of equipoise (DL or VL are required or contraindicated for patient safety).

- The primary outcome is sensible and explicitly defined as well.

Thank you for this feedback.

- The composite secondary outcome is interesting and captures an array of important patient-oriented outcomes. It would be helpful to report the four components separately as well as the composite incidences.

Thank you. In the results manuscript, we plan to report the secondary outcome both as a composite and by individual severe complications (hypoxemia, hypotension, cardiac arrest, and death).

- The definition of duration of laryngoscopy and tracheal intubation could be more robust. You state 'the final placement of an endotracheal tube or tracheostomy tube in the trachea'. Do you define the final placement as (a) insertion as declared by the intubator, (b) time to cuff inflation, (c) time to first capnography trace, or (d) other? It would be helpful to clarify this prior to data collection if possible, and if it is already defined, to have it more explicitly reported in the final manuscript.

In keeping with the pragmatic nature of our trial, training for time of final placement was operationalized by sites based on local protocols for confirming airway placement. Independent observers who collect the time of induction, laryngoscopy start, and final tube placement are trained by site PIs to collect time and attempt data the same way in both groups.

- The definition of oesophageal intubation is similarly contentious. Is it oesophageal intubation that is recognised visually or by virtue of an absent capnography trace or clinical signs? A clearer definition would be welcome as this is an important outcome associated with significant downstream morbidity when missed.

We agree with the reviewer that esophageal intubation is an important safety outcome. We have updated the manuscript to make it clear that esophageal intubation is reported by the operator immediately post-procedure as a clinical diagnosis based on the presence of any clinical sign, including visual inspection by the operator, absent capnography trace, absence of breath sounds, or absence of chest rise.

- The definitions of ICU-FDs and VFDs are clear and explicit.
- The remaining important outcomes are clearly defined.

Thank you for these comments.

- I note the distinct absence of overall intubation failure as a reported outcome. It would be extremely valuable to have this collected, especially in this high-risk cohort, as it is arguably a more robust and patient-oriented outcome than first-pass success. We used the definition of more than three failed attempts as failed intubation in our Cochrane review.

We appreciate this suggestion. We are using a granular definition of an intubation attempt, which is based on both laryngoscope insertion and tube/bougie insertion in the mouth. As the reviewer is aware, this is different than in previous trials. While we did not include failed intubation (> 3 unsuccessful attempts) as a pre-specified outcome in the initial trial protocol or trial registration on clinicaltrials.gov, we will consider including it as a post-hoc outcome in our final manuscript.

- The sample size calculation seems sensible. The absolute risk reduction of approximately 3% for failed intubation and approximately 5.5% for FPS (favouring VL) was noted in the recent Cochrane review. Provided that the remaining assumptions are realistic, it should result in a sufficiently powered study.
- The stopping boundary is appropriately conservative.

Thank you for these comments.

- I note the appropriate use of intention-to-treat analysis for the primary and secondary outcomes.

Given the significant potential for crossover, will a per protocol analysis be conducted and reported as well?

Thank you for this insightful question. We absolutely agree that failure to adhere to the randomly assigned laryngoscope type (a crossover event) potentially affects primary and secondary outcomes, as well as accurate interpretation of safety outcomes (Dodd S, et al. *Trials* 2012; 13:84). While recognizing that operators must prioritize patient safety during emergent procedures, we are attempting to minimize crossovers through trial training of site personnel and clinicians, actively monitoring adherence during the trial, and providing timely feedback to site principal investigators for operator re-education as needed. In our prior trials, the rates of crossover have been low enough to minimize the added value of per protocol analyses (e.g., in the BOUGIE trial, the rate of crossover was 1.4%). Thus, rather than specifying a per protocol analysis in this protocol, we plan a sensitivity analysis, in which any patient who experiences a crossover in study group assignment will be assumed to have failed to meet the primary outcome of successful intubation on the first attempt (page 22).

- I note the recommendations by the FDA and EMA for P value reporting, but do wonder whether it would make sense to report effect estimates with 95% CI for the primary and secondary outcome in the manuscript as well. This will aid future data extraction for evidence synthesis analyses.

We agree and thank the reviewer for this comment. We have revised the manuscript to be clear that, in addition to p values for the main analyses, we will provide effect sizes for the primary and secondary outcomes and 95% confidence intervals around those estimates.

- The set of variables collected is comprehensive and clearly defined.
- The secondary analyses are meaningfully defined and the predicted hypotheses in agreement with what is available in the literature to date.

Thank you for these comments.

General comments and observations:

- Consider replacing all instances of ‘endotracheal’ tube/intubation with ‘tracheal’ tube/intubation for uniformity with other published works on the topic. Orotracheal as a term, as used in the inclusion criteria, seems sensible as well.

We agree with the reviewer regarding the use of “tracheal intubation” and have used that term throughout the manuscript. We limit the use of “endotracheal” to uses of “endotracheal tube,” which is the preferred nomenclature in airway literature from the United States when referring to the tube specifically. We would prefer to keep that specific nomenclature unless the editor feels strongly otherwise.

Reviewer: 2

Comments to the Author:

In the manuscript, Prekker and co-authors present the study protocol for a prospective randomised trial of airway management. The hypothesis is that in the comparison of direct laryngoscopy vs. videolaryngoscopic assisted laryngoscopy differences in successful endotracheal intubation are to be achieved.

Adverse events are particularly associated with emergency airway management and show the highest risk of complications. In this context, hypoxia is one of the most common causes of airway management-associated deaths. In this respect, the primary endpoint for this study is correctly chosen.

Thank you for this perspective.

The increasingly common use of videolaryngoscopes leads to new and additional risks for airway injuries, as rigid guide rods are often used. Therefore, airway injuries should be included as secondary endpoints in the study.

Our case report form captures both injury to the teeth (a pre-specified exploratory safety outcome) and injury to airway structures. Data on both complications will be reported in the final manuscript.

The incidence of difficult mask ventilation is 1.5%, that of impossible mask ventilation 0.16%. The incidence of difficult direct laryngoscopy is 1.5% to 8.0% depending on the collective, the incidence of difficult intubation is slightly lower. An unexpected "cannot intubate, cannot ventilate" situation has a probability of 0.008% (1:13,000) to 0.004% (1:25,000). These incidences should also be recorded as they are significantly higher under emergency conditions.

We absolutely agree with the reviewer that difficult laryngoscopy and difficult tracheal intubation are important to distinguish in a trial of the effect of video laryngoscope use vs. direct laryngoscope use on first attempt intubation success. Our definition of successful intubation on the first attempt (the primary outcome) requires both 1 laryngoscopy attempt and 1 intubation attempt, as assessed by an independent observer, so we will be able to provide detailed data in the final manuscript on both difficult laryngoscopy and difficult intubation. Furthermore, we capture reasons for first attempt failure (inadequate view of the vocal cords, difficulty passing the tube, and difficulty passing the bougie are all included). Unexpected “cannot intubate, cannot ventilate” scenarios are exceedingly rare, as the reviewer notes, so we do not capture this occurrence per se in the case report form. We do record cricothyrotomy procedures, perhaps a reasonable surrogate.

Hyperangulated spatulas are used. Can we assume that the video laryngoscope with hyperangulated blades are of the D-BLADE type? Or are different blades used?
Since the study is multicentre, this is an important point.

Thank you for this point of clarification. Reviewer #1 asked something very similar. We record the brand and size of the hyperangulated laryngoscope used for the first intubation attempt and will report these data in the final manuscript. Trial sites have one or more of the following hyperangulated video laryngoscopes available: Storz D Blade, GlideScope G/AVL, and GlideScope LoPro.

I doubt that, as the authors describe on page 20, it makes no difference whether intubation is done in the ED or in the ICU. The patient populations are very different, the risk of aspiration is higher, the oxygen reserve in the ICU is lower (as the indication is more likely to be given later), and so on.

We agree with the reviewer that characteristics and outcomes of patients undergoing emergent tracheal intubation differ between the ED and the ICU. However, our hypothesis on page 20 (under ‘Effect Modification’) states that patient location will not modify the effect of VL vs. DL on the occurrence of successful intubation on the first attempt. One could speculate that the overall proportion of patients experiencing successful intubation on the first attempt may be lower in the ICU compared to the ED, for reasons including those highlighted by the reviewer, but we are hypothesizing that the effect of VL vs. DL (benefit/no benefit/harm) will be similar in both settings (i.e., no evidence for heterogeneity of treatment effect).