Comparison of a Novel Cadaver Model (Fix for Life) With the Formalin-Fixed Cadaver and Manikin Model for Suitability and Realism in Airway Management Training

Michael W. van Emden, MD,* Jeroen J. Geurts, PhD,* Patrick Schober, MD, PhD† and Lothar A. Schwarte, MD, PhD†

BACKGROUND: Manikins are widely used in airway management training; however, simulation of realism and interpatient variability remains a challenge. We investigated whether cadavers embalmed with the novel Fix for Life (F4L) embalmment method are a suitable and realistic model for teaching 3 basic airway skills: facemask ventilation, tracheal intubation, and laryngeal mask insertion compared to a manikin (SimMan 3G) and formalin-fixed cadavers.

METHODS: Thirty anesthesiologists and experienced residents (“operators”) were instructed to perform the 3 airway techniques in 10 F4L, 10 formalin-fixed cadavers, and 1 manikin. The order of the model type was randomized per operator. Primary outcomes were the operators’ median total ranking of each model type as a teaching model (total rank), ranking of the model types per technique, and an operator’s average verbal rating score for suitability and realism of learning the technique on the model. Secondary outcomes were the percentages of successfully performed procedures per technique and per model (success rates in completing the respective airway maneuvers). For each of the airway techniques, the Friedman analysis of variance was used to compare the 3 models on mean operator ranking and mean verbal rating scores.

RESULTS: Twenty-seven of 30 operators (90%) performed all airway techniques on all of the available models, whereas 3 operators performed the majority but not all of the airway maneuvers on all models for logistical reasons. The total number of attempts for each technique was 30 on the manikin, 292 in the F4L, and 282 on the formalin-fixed cadavers. The operators’ median total ranking of each model type as a teaching model was 1 for F4L, 2 for the manikin and, 3 for the formalin-fixed cadavers (P < .001). F4L was considered the best model for mask ventilation (P = .029) and had a higher mean verbal rating score for realism in laryngeal mask airway insertion (P = .043). The F4L and manikin did not differ significantly in other scores for suitability and realism. The formalin-fixed cadaver was ranked last and received lowest scores in all procedures (all P < .001). Success rates of the procedures were highest in the manikin.

CONCLUSIONS: F4L cadavers were ranked highest for mask ventilation and were considered the most realistic model for training laryngeal mask insertion. Formalin-fixed cadavers are inappropriate for airway management training. (Anesth Analg 2018;127:914–9)

KEY POINTS

• Question: Is a novel cadaver model a suitable and realistic model for acquiring airway management techniques?
• Finding: The novel cadaver model was ranked best for mask ventilation and was considered the most realistic for laryngeal mask insertion.
• Meaning: The novel cadaver model could be a suited and more realistic alternative to manikins in training airway management techniques to novices.
without prior consent is not to be advocated.9 The use of formalin-fixed cadavers of body donors has no time constraints, and such cadavers can be used for many years for the teaching of medical students and for scientific purposes. However, toxic levels of formaldehyde and tissue rigidity are a matter of concern.10,11 Recently, the embalming method Fix for Life (F4L) has been developed, which allegedly preserves the haptic quality of fresh human material while only minimal amounts of formaldehyde are needed.12 A wide scope of opportunities for teaching and learning airway management techniques as well as other procedures in a controlled setting could be possible with this embalming method without the disadvantages of fresh material. The aim of our study was to determine whether the F4L cadavers are a suited and realistic airway management teaching model in comparison with an established simulation manikin and formalin-fixed cadavers.

METHODS
The Medical Ethics Review Committee of VU University Medical Center, Amsterdam, the Netherlands (VUmc), judged the Medical Research Involving Human Subjects Act not to be applicable, and official approval for the study was thus not required in accordance with Dutch law. This manuscript adheres to the applicable Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.13 The feasibility of 3 different models to train several airway techniques (facemask ventilation, tracheal intubation, and laryngeal mask insertion) was studied. Given the anatomical heterogeneity of the human body compared to the homogeneity of the manikin, 10 F4L and 10 formalin-fixed cadavers were used and compared to 1 manikin to investigate whether the F4L-embalmed cadavers are a suited teaching model for novices to acquire airway management skills. Experienced airway management providers were asked to rate suitability and realism of the 3 models as detailed below.

Models
The cadavers were all donated to the Department of Anatomy and Neurosciences, VUmc. In accordance with Dutch law, all body donors had given written consent for body donation to science after death. The embalming of the F4L- and formalin-fixed cadavers was performed in the Department of Anatomy and Neurosciences, VUmc. The embalming commenced within 24–72 hours after death. Infusion of the formalin embalment fluid (a mixture of formaldehyde, alcohol, salicylic acid, various salts, thymol, and water) and the F4L embalment fluid (a patented mixture of aldehyde, other nonhazardous components and a small amount of formaldehyde) was via the femoral artery. Further embalment procedures were done according to the prescribed method for a formaldehyde or F4L fixation. The manikin used was the SimMan 3G (Laerdal Medical, Stavanger, Norway), an established high-fidelity model for acquiring airway management skills and research in the simulation setting.3,14–16

Experimental Protocol
Thirty physician anesthesiologists and senior residents in their third to fifth year of the 5-year residency training program were approached to perform the several airway techniques on the cadavers and manikin after obtaining informed consent. The inclusion criterion for participation was at least 100 successful tracheal intubations. Exclusion criteria were pregnancy or lactation of the operator because of teratogenicity of formaldehyde.10,11 Each operator was randomized by a sealed envelope technique for the order of the 3 model types in which the airway techniques were performed. The operator was instructed to complete all airway techniques on all the cadavers per model type before continuing to the next model type. All operators performed the techniques individually and were instructed not to discuss their experience with other operators. All the data were collected in the dissection room of the Department of Anatomy and Neurosciences, VUmc between December 2016 and February 2017.

All models were placed in the supine position, and positioning could be optimized (eg, sniffing position) by the individual operator at his or her discretion using cushions. Any remaining oral fluids in the cadavers were suctioned. Each operator was asked to perform mask ventilation first, followed by tracheal intubation and finally insertion of a laryngeal mask (LM) per model. A successful procedure was defined as visible chest movements by ventilating with a self-inflating bag resuscitator (Manual Resuscitator Adult, Hsiner Co, Ltd, Taichung City, Taiwan) within 30 seconds for mask ventilation and within 2 minutes for the other 2 techniques. Chest movements are difficult to detect in formalin-fixed cadavers due to tissue rigidity. These models were dissected, allowing direct assessment of airflow through the trachea. For mask ventilation, adult face masks sizes 4 and 5 (Air Cushion Mask with Valve, Hsiner Co, Ltd, Taichung City, Taiwan) and Guedel airways sizes 5 and 6 (Mallinckrodt DAR S.r.l., Modena, Italy) were available at the operators’ discretion. For the tracheal intubation, Macintosh laryngoscope blade sizes 3 and 4 (EmdaMed, Berkel en Roderijs, the Netherlands) as well as an intubating catheter (Frova, William Cook Europe ApS, Bjaeverskov, Denmark) could be used. Tracheal tubes were available in sizes 7.0 and 8.0 mm (Mallinckrodt Hi-Contour Oral/Nasal Tracheal Tube Cuffed, Coviden Ibc, Mansfield, MI). For LM airway insertion, the available laryngeal masks were sizes 4 and 5 (PROACT Medical Ltd, Corby, UK). The use of water was allowed as a lubricant. During each airway management procedure, one assistant aided the operator on request (eg, to hand equipment or to perform backward, upward, or rightward pressure [BURP] of the larynx),17 and another documented the data on a standardized data form. Mask ventilations were objectively classified with the Han score, in which grade 1 defines ventilation by mask without aids, grade 2 describes ventilation by mask with oral airway (Guedel), grade 3 defines difficult mask ventilation requiring 2 practitioners, and grade 4 indicates inability to mask ventilate.18 For tracheal intubation, the Cormack-Lehane grades were documented as reported by the operators.19 After each airway management procedure, the operator was asked to give a verbal rating score (VRS) for suitability (defined as the operator’s assessment of suitability of the model in teaching the novice the performed airway management skill) and for realism (defined as the operators assessment of look, feel and flexibility of the model compared to real patients) of the model, both on a scale of 1–10 (1 = worst score to 10 = best score).

Specifically for the F4L cadavers, the operators were additionally asked for VRSs for suitability and realism of the
model in learning to manage the difficult airway. After completing all airway techniques on all of the different models, the operators were instructed to consider all aspects of suitability and realism of the models in teaching the novice airway management skills and to rank the models accordingly. The operators were asked to rank the models first, second, and third separately for mask ventilation, tracheal intubation, and LM placement. Finally, the operators gave an overall total rank from 1 to 3 to the different models with respect to all aspects of suitability and realism in airway management skills, considering all the airway techniques together.

The primary outcome measures were the total ranking of the different models as airway teaching model, the ranking of the models per technique, and the VRSs for suitability and realism of the models per technique. Secondary outcomes were success rates of the different techniques per model.

**Statistical Analysis**

IBM SPSS Statistics for Windows, Version 22.0. (IBM Corp, Armonk, NY) was used for statistical analysis. The Friedman analysis of variance (ANOVA) was used to test the null hypothesis that ranking scores and VRSs are equally distributed among the F4L model, the formalin-fixed cadaver model, and the manikin model. When significant, pairwise comparisons were performed, and P values were corrected for multiple comparisons using the Dunn-Bonferroni test. Two-sided P values <.05 were considered significant.

While the Friedman ANOVA accounts for correlations between repeated measurements across the 3 model types, it does not account for within-participant correlation of repeated assessments of VRSs within each model type. Because we were interested in an overall VRS per technique per model rather than in individual scores for each cadaver, we reduced each individual’s sequence of measurements within each model to a single number (ie, the mean of the score per technique and model). This “summary statistic approach,” as recommended by Senn et al20 and Matthews et al,21 eliminated within-participant correlation within model, allowing a comparison of VRSs between the models with the Friedman ANOVA.

**Sample Size**

Sample size estimations were performed with STATA 13.0 (STATACorp, College Station, TX). We planned to use nonparametric tests for the analysis of ranks and ordinal rating scales, for which exact calculation of sample size is not possible. Therefore, we calculated the sample size for the parametric equivalent (here: repeated-measures ANOVA) and added 15% to this number as compensation.22 We aimed to choose a sample size such that a 2-point score difference in verbal rating scale between any 2 models could be detected with 90% power on a .05 α level. Previous data on the correlation between repeated measurements of different participants in the 3 model types and estimates of the error variance were not available. We therefore performed sample size calculations using different correlations (0.1–0.9) and different error variances (0.5–5). The largest calculated sample size was 23 participants. Adding 15% for the nonparametric test, we would need 27 participants. To account for possible drop out of participants, we targeted 30 participants.

**RESULTS**

In both the F4L and formalin-fixed cadaver group, there were 7 male and 3 female cadavers. Mean age of demise was similar: 79 years in the F4L group, and 80 years in the formalin-fixed group. Dental status was also similar in both groups; 5 F4L and 6 formalin-fixed cadavers had no teeth. The body habitus in both cadaver groups was also comparable, and cadavers with extremes in body composition (eg, extreme obesity or cachexia) were not present in either group. Finally, mean thyromental distance was around 6.5 cm in both cadaver groups (6.3 cm in F4L, 6.7 cm in formalin) and 7.0 cm in the manikin group.

Of the operators, 14 were females, and 16 were males. Mean age was 38.5 years (range 28–56 years). Twenty were anesthesiologists, and 10 were senior residents, with a mean experience in airway management of 10.3 years (range 3–33 years). Twenty-seven of 30 operators (90%) performed all airway techniques on all of the available models (ie, 21 mask ventilations; 21 LM airway insertions; and 21 intubations in 10 formalin-fixed cadavers, 10 F4L cadavers, and 1 manikin), whereas 3 operators performed the majority but not all of the airway maneuvers for logistical reasons. However, all participants did perform all airway techniques in each of the model types. The total number of attempts for each technique was 30 on the manikin, 292 in the F4L, and 282 on the formalin-fixed cadavers.

**Ranking Outcomes**

Median total ranking scores as airway training model for the F4L, manikin, and formalin-fixed cadaver models were 1, 2, and 3, respectively (P < .001). In the ranking of mask ventilation, F4L was ranked significantly higher compared to the manikin and formalin-fixed cadaver. For intubation and LM airway insertion, both F4L and manikin ranked higher than the formalin-fixed cadaver. The other ranking outcomes for mask ventilation, tracheal intubation, and LM airway insertion are summarized in Table 1.

**Verbal Rating Scores**

The mean VRSs regarding suitability and realism of the different models in teaching skills to novice providers for the different airway management techniques are presented in Table 2. The mean VRSs were overall significantly different (P < .001). Pairwise comparisons showed that F4L had a significantly higher VRS for realism in LM airway insertion compared to the manikin (P = .043). All the VRSs of the F4L or manikin compared to the formalin-fixed cadavers were significantly higher (P < .001). The mean VRSs for suitability and realism of the F4L cadaver as a difficult airway teaching model were 7.7 of 10 (95% CI, 7.2–8.2) and 7.0 of 10 (95% CI, 6.4–7.6), respectively.

**Success Rates**

In mask ventilation, 100% of attempts were successful on the manikin, 95.9% in F4L, and 96.8% in formalin-fixed cadavers. The success rates in intubation were 100% in manikins, 61% in F4L, and 0% in formalin-fixed cadavers. For LM airway insertion, 73.3% of attempts in the manikin, 69.5% in F4L, and 2.8% in formalin-fixed cadavers were successful. Han scores and Cormack-Lehane grades, application of BURP, and use of an intubating catheter in the different models are presented in the Figure.
DISCUSSION

This is the first study to use F4L-embalmed cadavers as an airway management training model. We investigated the suitability and realism of this cadaver model compared to an established airway training manikin and standard formalin-fixed cadavers to teach essential airway management skills. Our study demonstrates that the F4L cadaver scores better than the manikin in some aspects of teaching these procedures, while it is as good as the manikin regarding other aspects. More specifically, in the overall score for mask ventilation, the F4L cadaver was ranked significantly higher than the manikin. Moreover, the F4L cadaver was considered a more realistic model for the insertion of an LM airway. In addition, the F4L cadaver received promising VRSs regarding suitability and realism for teaching difficult airway management.

Although simulation training has found a permanent place in daily anesthesiology practice and training programs and trainees seem to benefit from it (eg, in training complex team scenarios), there is an ongoing discussion about the...

Table 1. Summary of Ranking Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Median Rank (Range)</th>
<th>Pairwise Comparisons (P Values)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F4L (n = 30)</td>
<td>Manikin (n = 30)</td>
</tr>
<tr>
<td>Total ranking</td>
<td>1 (1–1)</td>
<td>2 (1–2)</td>
</tr>
<tr>
<td>Mask ventilation</td>
<td>1 (1–2)</td>
<td>2 (1–3)</td>
</tr>
<tr>
<td>Intubation</td>
<td>2 (1–2)</td>
<td>1 (1–2)</td>
</tr>
<tr>
<td>LM airway insertion</td>
<td>1 (1–2)</td>
<td>2 (1–2)</td>
</tr>
</tbody>
</table>

Median for total ranking of the models and median rank per technique (range in brackets). Ranking: 1 = best, 3 = worst. P values are given for pairwise comparisons performed after significant Friedman analysis of variance. A P value < .05 after correcting for multiple comparisons was considered a significant difference in ranking.

Abbreviations: F4L, Fix for Life; LM, laryngeal mask.

Table 2. Mean Verbal Rating Scores

<table>
<thead>
<tr>
<th>Technique</th>
<th>Mean Score (95% CI)</th>
<th>F4L</th>
<th>Manikin</th>
<th>Formalin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mask ventilation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suitability</td>
<td>7.4 (7.0–7.8)</td>
<td>7.2 (6.6–7.7)</td>
<td>3.4 (2.7–4.1)*</td>
<td></td>
</tr>
<tr>
<td>Realism</td>
<td>5.9 (5.3–6.6)</td>
<td>7.0 (6.5–7.5)</td>
<td>2.8 (2.2–3.4)*</td>
<td></td>
</tr>
<tr>
<td>Intubation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suitability</td>
<td>7.2 (6.7–7.7)</td>
<td>5.4 (4.8–6.0)</td>
<td>1.3 (1.1–1.4)*</td>
<td></td>
</tr>
<tr>
<td>Realism</td>
<td>5.9 (5.3–6.6)</td>
<td>5.7 (5.1–6.2)</td>
<td>1.6 (1.2–2.0)*</td>
<td></td>
</tr>
<tr>
<td>LM airway insertion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suitability</td>
<td>4.2 (3.4–5.0)</td>
<td>5.0 (4.5–5.5)</td>
<td>1.2 (1.1–1.3)*</td>
<td></td>
</tr>
<tr>
<td>Realism</td>
<td>3.6 (2.8–4.5)</td>
<td>5.0 (4.4–5.6)*</td>
<td>1.6 (1.2–1.9)*</td>
<td></td>
</tr>
</tbody>
</table>

Mean scores with 95% CI for suitability and realism per model in teaching the different techniques. For the F4L and formalin-fixed cadaver models, the data being summarized are the mean scores across the multiple observations per operator, while a single score per operator was obtained for the manikin. A P value < .05 after correcting for multiple comparisons was considered a significant difference in ranking.

Abbreviations: CI, confidence interval; F4L, Fix for Life; LM, laryngeal mask.

*Significant pairwise comparison (P < .001) versus all other models.

**Significant pairwise comparison (P = .043) versus the manikin.

Figure.

- Percentages of Han scores (A), Cormack-Lehane grades (B), use of backward upward rightward pressure (BURP) and intubating catheter (C) in the different models. Low percentages of BURP and use of an intubating catheter in formalin-fixed cadavers resulting from operators not attempting laryngoscopy because of the rigidity of the jaw. F4L indicates Fix for Life.

VRGs regarding suitability and realism for teaching difficult airway management.

Although simulation training has found a permanent place in daily anesthesiology practice and training programs and trainees seem to benefit from it (eg, in training complex team scenarios), there is an ongoing discussion about the...
validity and reliability of airway skill simulators). As an alternative to manikin simulators, cadavers can be used to train airway management techniques. Fresh frozen cadaver models have been found useful to train airway management techniques. However, the logistics required for defrosting and the limited timespan in which the material is available for use because of decomposition limits the usefulness as teaching model. Only 1 previous study assessed embalmed cadavers as an airway management training model. The study by Szücs et al. used the Thiel embalmment method for cadavers, which has also been used and studied in other contexts such as surgical procedures. Similar to our data, these authors found higher mean VRSs for mask ventilation for the cadaver model compared to manikins. However, while Thiel-embalmed cadavers need to be stored for several months for optimal embalmment, F4L-embalmed cadavers are immediately available for use and can be used for a minimum of 2 years (presumably much longer, but long-term experience is not yet available because the technique is relatively new). Our findings suggest that the F4L embalmment method can be a useful alternative to the Thiel method to train airway management techniques in anatomical institutions.

This present study is the first to show that formalin-fixed cadavers do not seem suited for the acquisition of airway management skills, as shown by the low ranking and verbal rating scales. A likely explanation is the rigidity of the cadaver resulting from embalming with formaldehyde.

One recent study analyzed the biomechanical properties of human spines embalmed with F4L and found an increase in spinal tissue stiffness. This finding might explain our results of a relatively high percentage of Cormack–Lehane grades 3 and 4, application of BURP, and insertion of an intubating catheter in the F4L cadaver. These results might also explain the high VRSs given in the assessment of suitability and realism (7.7 and 7.0, respectively) of the F4L cadaver as a model for difficult airway procedures. Possible future adjustments to the F4L mixture in relation to the amount of formaldehyde used could lead to even more flexibility of the joints of the F4L cadaver.

There are limitations to our study. First, it was not possible to blind the participants with respect to the model type. Second, there is no universally accepted or validated method to measure suitability or realism of an airway model. Because we needed to rely on subjective assessments, we deliberately recruited participants with extensive experience. Literature suggests that approximately 50 tracheal intubations are required to achieve a plateau phase in this skill. We chose to double this number and defined 100 successful tracheal intubations as the minimum requirement. This ensured that all participants were well experienced, such that they all had a common framework to which they could compare the realism and suitability of the models. The participants’ assessment of suitability and realism was measured using a verbal rating scale. In a previous study by Szücs et al., a similar verbal rating scale was used and demonstrated that this scale is useful to reveal differences between airway models. Moreover, similar scales are abundantly used in other fields of science for quantification of subjective assessment, and these scores are well validated for this purpose. Examples include the subjective assessment of pain and numerous scales in psychology.

Third, we only used one type of manikin, while there are various manikins available on the market, and our results may not be applicable to other manikins. However, we selected a well-established and widely used manikin particularly for airway management. Our observed success rates of intubation and LM airway insertion are comparable to those reported for the SimMan by Schebesta et al. (97.5% and 67.5%, respectively). We only used 1 manikin compared to 10 F4L and 10 formalin-fixed cadavers. This is because manikins of one type are virtually identical, while human bodies differ in anatomy from individual to individual, which approaches daily practice more closely. Older age, presence of a beard, and lack of teeth are associated with a more difficult mask ventilation. The cadavers studied had a mean age of 80 years, and some lacked dentition. This could explain why in the F4L cadavers, Han scores 3 (requiring 2 practitioners) and 4 (impossible mask ventilation) were reported more often than in manikins.

Fourth, we provided a limited range of airway instruments the operators could use. However, the adult facemask, a standard Macintosh laryngoscope blade, and an LM are the basic airway instruments, and a novice airway practitioner should be acquainted with these early in his or her career. More advanced airway instruments and techniques such as the video laryngoscope or fiber optic intubation should be studied in the F4L cadaver in the future.

In conclusion, we found that the F4L cadaver model was judged by specialists to be the most appropriate model for teaching mask ventilation, as well as the most realistic model for LM airway insertion. We see the potential for the F4L cadaver in skills training to manage the more difficult airway and possibly other acute medical procedures. The formalin-fixed cadavers are inappropriate for airway management training.

ACKNOWLEDGMENTS

The authors thank Eliane Kaaij and Jasmina Rubira Yoxall, dissection room staff, Department of Anatomy and Neurosciences, VU University Medical Center, Amsterdam, the Netherlands, who helped with collecting the data.

DISCLOSURES

Name: Michael W. van Emden, MD.
Contribution: This author helped design the study, conduct the study, analyze the data, and write and revise the paper.

Name: Jeroen J. Geurts, PhD.
Contribution: This author helped design the study, conduct the study, analyze the data, and write and revise the paper.

DISCLOSURES

Name: Patrick Schober, MD, PhD.
Contribution: This author helped design the study, conduct the study, analyze the data, and write and revise the paper.

Name: Lothar A. Schwarte, MD, PhD.
Contribution: This author helped design the study, conduct the study, analyze the data, and write and revise the paper.

This manuscript was handled by: Richard P. Dutton, MD.

REFERENCES
