Anaesthesia 2021, 76, 366–372 doi:10.1111/anae.15238

Original Article

Anatomic accuracy of airway training manikins compared with humans*

M. B. Blackburn,[1] **S. C. Wang,**[2] **B. E. Ross,**[3] **S. A. Holcombe,**[4] **K. M. Kempski,**[5] **A. N. Blackburn,**[6] **R. A. DeLorenzo**[7] and **K. L. Ryan**[8]

1 Research Physiologist, 8 Capability Area Manager, United States Institute of Surgical Research, Joint Base San Antonio, Fort Sam, Houston, TX, USA
2 Executive Director and Professor, 3 Research Analyst, 4 Assistant Research Professor, Morphomic Analysis Group, University of Michigan, Ann Arbor, MI, USA
5 Research Fellow, Johns Hopkins University, Baltimore, MD, USA
6 Visiting Assistant Professor, St. Mary's University, San Antonio, TX, USA
7 Professor, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

Summary

Airway simulators, or training manikins, are frequently used in research studies for device development and training purposes. This study was designed to determine the anatomic accuracy of the most frequently used low-fidelity airway training manikins. Computerised tomography scans and ruler measurements were taken of the SynDaver®, Laerdal® and AirSim® manikins. These measurements were compared with human computerised tomography (CT) scans (n = 33) from patients at the University of Michigan Medical Center or previously published values. Manikin measurements were scored as a percentile among the distribution of the same measurements in the human population and 10 out of 27 manikin measurements (nine measurements each in three manikins) were outside of two standard deviations from the mean in the participants. All three manikins were visually identifiable as outliers when plotting the first two dimensions from multidimensional scaling. In particular, the airway space between the epiglottis and posterior pharyngeal wall, through which airway devices must pass, was too large in all three manikins. SynDaver, Laerdal and AirSim manikins do not have anatomically correct static dimensions in relation to humans and these inaccuracies may lead to imprecise airway device development, negatively affect training and cause over-confidence in users.

Correspondence to: M. B. Blackburn
Email: megan.b.blackburn2.civ@mail.mil

Accepted: 26 July 2020
Keywords: airway; anatomic accuracy; manikin; trainers

*Presented in part at the Military Health Systems Research Symposium, Kissimmee, FL, USA, August 2019.

Twitter: @AugustBlackburn

Introduction

Tracheal intubation can be a critical life-saving intervention and around 25 million are performed in the USA per year in a variety of patients, from those undergoing elective surgery in the operating theatre, to those who require mechanical ventilation due to cardiorespiratory illness or trauma. It is a technically challenging skill to learn and maintain [1], requiring up to 75 iterations to reach proficiency [2], and is associated with a high complication and failure rate. Data suggest that as the number of attempts to secure the airway increases, so does the incidence of adverse events [3]. Airway compromise, for example, is the second leading cause of potentially survivable death on the battlefield, accounting for 1 in 10 preventable deaths [4]. Additionally, in the civilian setting, success of tracheal intubation depends largely on training level with emergency medical technicians, paramedics and physicians having up to 50%, 10–20% and 10–15% failure rates on first-pass attempts respectively [5–9]. This supports the notion that airway management is highly dependent on skill and training.

Manikins are used for a variety of purposes in the airway management field given the practical and ethical limitations of training on live humans, and have been in use since the 1960s when Laerdal® developed the first airway training manikin for mouth to mouth resuscitation [10]. Some of the earliest simulators produced were for testing patient safety under anaesthesia by allowing reproduction of high-risk situations without risk to human subjects, similar in concept to training pilots. There are now simulators of various fidelity used for training prehospital providers [11], testing prototype devices and technologies [12], and assessing provider capabilities [13, 14]. Given such extensive use, it is imperative that findings be put into context and manikins be verified as fit for purpose.

Only a few studies have examined the anatomical correctness of airway manikins with a specific focus on high fidelity trainers [15, 16]. To our knowledge, there are very few investigating the anatomical correctness of the lower cost, low-fidelity manikins, which are those most commonly used for device development and scientific experiments. Furthermore, there are no studies that include the more recently developed SynDaver® technology, which incorporates synthetic body parts and aims to 'replace live animals, cadavers, and real patients in medical device evaluation, clinical training, and medical simulation' (https://syndaver.com). For the purpose of our study, we examined upper airway spaces through which an airway device would pass. We hypothesised that all airway manikin structures would allow for less complicated tracheal intubation than those found in the human population due to inaccurate anatomic structures.

Methods

This study was approved by the University of Michigan Institutional Review Board. Patients who were cared for at the University of Michigan between 2009 and 2017 were studied retrospectively. Thirty-three participants aged between 18 and 47 years (15 female, 18 male, 3 African American, 1 Asian, 25 Caucasian, 4 other) with BMIs between 20.9 and 40.6 kg.m^{-2} who underwent head and neck computerised tomography (CT) angiogram were selected. The angiograms were acquired on General Electric LightSpeed® (GE Healthcare, Chicago, IL, USA) machines using 100 kVp for individuals < 30 kg.m^2 BMI or 120 kVp for individuals ≥ 30 g.m^2 BMI, modulated current between 150 and 625 mA, 0.625 mm slice thickness and interval, using a standard reconstruction kernel from the aortic arch to cranial vertex. Distance measurements of the upper airway were done using Mimics (version 18, Materialise NV, Leuven, Belgium) from the helical reconstruction.

Upper airway CT images were taken of three low-fidelity airway trainer manikins: SynDaver (Standard Adult Airway Trainer, SynDaver Labs, Tampa, FL, USA), Laerdal (Airway Management Trainer, Laerdal Medical, Stavanger, Norway) and AirSim® (AirSim Advance Model, TruCorp, Belfast, Ireland). Computerised tomography images included the nasal cavity to the main stem bronchi (Toshiba Aquilion Prime 160, 120 kV, 50 mAs, 2 mm slice thickness, 0.637 pitch factor, 0.5 × 80 collimation). The manikins were in neutral head position during scanning. Static dimensions and angles of the upper airway from CT scans were measured using VitreaCore (Version 6.7.4; Vital Images, Minnetonka, MN, USA). Following CT scans, manikins were cut down the sagittal midline using a band saw and additional measurements were taken manually using a ruler and goniometer. All distances were taken in the sagittal mid-plane three times and averaged in order to optimise measurement accuracy.

Computerised tomography measurements included: (1) horizontal distance from the posterior edge of the tongue to the posterior pharyngeal wall; (2) horizontal distance from the tip of the epiglottis to the posterior pharyngeal wall; (3) linear distance from the tip of the tongue to the vallecula; (4) linear distance perpendicular from the midpoint on the line connecting the tip of the tongue and the vallecula to the tongue dorsum; (5) linear distance from the vallecula to the tip of the epiglottis; (6) horizontal distance from the base of the epiglottis to the posterior pharyngeal wall; (7) vertical distance of the soft palate; (8) vertical distance from the base of the soft palate to the laryngeal inlet; and (9) vertical distance from the tip of the soft palate (uvula) to the tip of the epiglottis (Fig. 1a and b). Additional ruler measurements included: (10) height mouth opens; (11) first tracheal ring width; (12) tracheal length; and (13) neck circumference (Fig. 1c).

For the comparison of measurements using CT, the distribution of anatomical measurements among participants was modelled as a normal distribution and summarised using the mean and standard deviation. Shapiro-Wilk normality tests for each of the measurements were not statistically significant, indicating the assumption of normality is reasonable. We also compared manikins with previously published measurements, which were reported

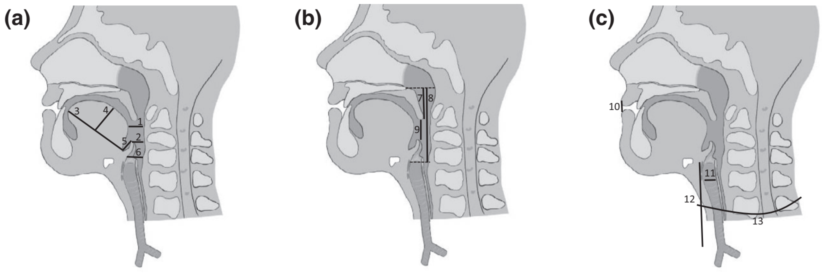

Figure 1 Airway CT measurements: (a-1) horizontal distance from the posterior edge of the tongue to the posterior pharyngeal wall (PPW); (2) horizontal distance from the tip of the epiglottis to the PPW; (3) linear distance from the tip of the tongue to the vallecula; (4) linear distance perpendicular from the midpoint on the line connecting the tip of the tongue and the vallecula to the tongue dorsum; (5) linear distance from the vallecula to the tip of the epiglottis; (6) horizontal distance from the base of the epiglottis to the PPW and (b-7) vertical distance of the soft palate; (8) vertical distance from the base of the soft palate to the laryngeal inlet; (9) and vertical distance from the tip of the soft palate (uvula) to the tip of the epiglottis. Airway ruler measures: (c-10) height mouth opens; (11) first tracheal ring width; (12) trachea length; and (13) neck circumference.

to be or assumed to be normally distributed [16, 17–20]. Percentile scores were calculated for the manikins based on a normal distribution defined by parameters calculated from the participants or defined in previous reports. Visual interpretation of the representativeness of manikins was performed by plotting the first two dimensions from classical multidimensional scaling; a non-linear dimensionality reduction technique commonly used for viewing the similarity/dissimilarity between objects (in this case humans and manikins) that have high-dimensional data. Here, the matrix of dissimilarities is composed of Euclidean distances calculated using the CT data [21, 22] from human participants and the three manikins.

Results

Thirty-three participants were included in this study. Mean (SD) age was 33.7 (8.2) years and BMI 27.5 (5.4) kg.m^{-2}. Table 1 shows the percentile that manikin measurements fall within the distribution of participant measurements. A percentile of 0.5 represents an anatomically accurate manikin while percentiles near 0 or 1 indicate anatomically incorrect measurements compared with participants. Ten of 27 manikin measurements (nine measurements each in three manikins) were outside of two standard deviations from the mean in the participants. SynDaver, Laerdal and AirSim had three, five and two measures outside of two standard deviations respectively. Some common trends across manikins were apparent. Three measurements were larger than the mean plus one standard deviation in participants for all three manikins: the distance from the tongue to the posterior pharyngeal wall, the distance from the tip of the epiglottis to the posterior pharyngeal wall, and the distance from the base of the epiglottis to the posterior pharyngeal wall. Additionally, all three measurements for

Table 1 Summary of participant measures, manikin measures and manikin percentile scores assuming a normal distribution for human participants. Anatomically accurate manikins would have a percentile score of 0.5.

Measurement	Mean (SD) Human participants	Value (percentile) SynDaver	Laerdal	AirSim
Tongue to PPW	12.22 (5.42)	19.9 (0.96[a])	24.0 (> 0.99[a])	16.8 (0.86[a])
Epiglottis to PPW	7.94 (3.35)	15.4 (0.99[b])	23.5 (> 0.99[b])	11.3 (0.84[a])
Tip of tongue to vallecula	71.49 (6.01)	50.7 (< 0.01[b])	73.6 (0.64)	63.9 (0.1[a])
Tip of tongue to tongue dorsum	34.38 (5.25)	29.7 (0.18)	23.2 (0.02[b])	28.7 (0.14[a])
Vallecula to epiglottis	14.64 (4.2)	16.0 (0.62)	16.3 (0.66)	8.7 (0.08[a])
Base of epiglottis to PPW	11.84 (3.1)	23.9 (> 0.99[b])	28.7 (> 0.99[b])	16.0 (0.91[a])
Vertical distance of soft palate	26.50 (7.71)	15.5 (0.08[a])	41.3 (0.97[a])	11.1 (0.02[b])
Soft palate to laryngeal inlet	60.64 (9.97)	66.2 (0.71)	112.0 (> 0.99[b])	82.8 (0.99[b])
Uvula to epiglottis	21.40 (7.88)	28.1 (0.8)	28.4 (0.81)	25.0 (0.68)

PPW, posterior pharyngeal wall.
[a] Outside one SD.
[†] Outside two SD.

the vertical distance of the soft palate were outside of one standard deviation from the mean for all three manikins, but not in a consistent direction.

Figure 2 shows distributions for individual measurements. Histograms for each of the nine measurements are overlaid with the respective measure for all three manikins. Of particular importance are those manikin measurements that were > 99% or < 0.01%, such as tongue to posterior pharyngeal wall, epiglottis to posterior pharyngeal wall, tip of tongue to vallecula, base of epiglottis to posterior pharyngeal wall, and soft palate to laryngeal inlet. In order to consider the measures jointly, we plotted the first two dimensions from multidimensional scaling. In this scenario, an anatomically accurate manikin would be positioned near the centre of the cluster of participants. All three manikins were visual outliers among the participants as shown in Figure 3.

In order to capture measures that could not be taken via CT imaging analysis and ensure our comparisons held true across multiple populations, we compared manikins with previously published airway values (Table 2). With the exception of a few measurements, manikins were not representative of the central tendency in the human populations referenced [16, 17–20], as evidenced by percentiles near 0 and 1.

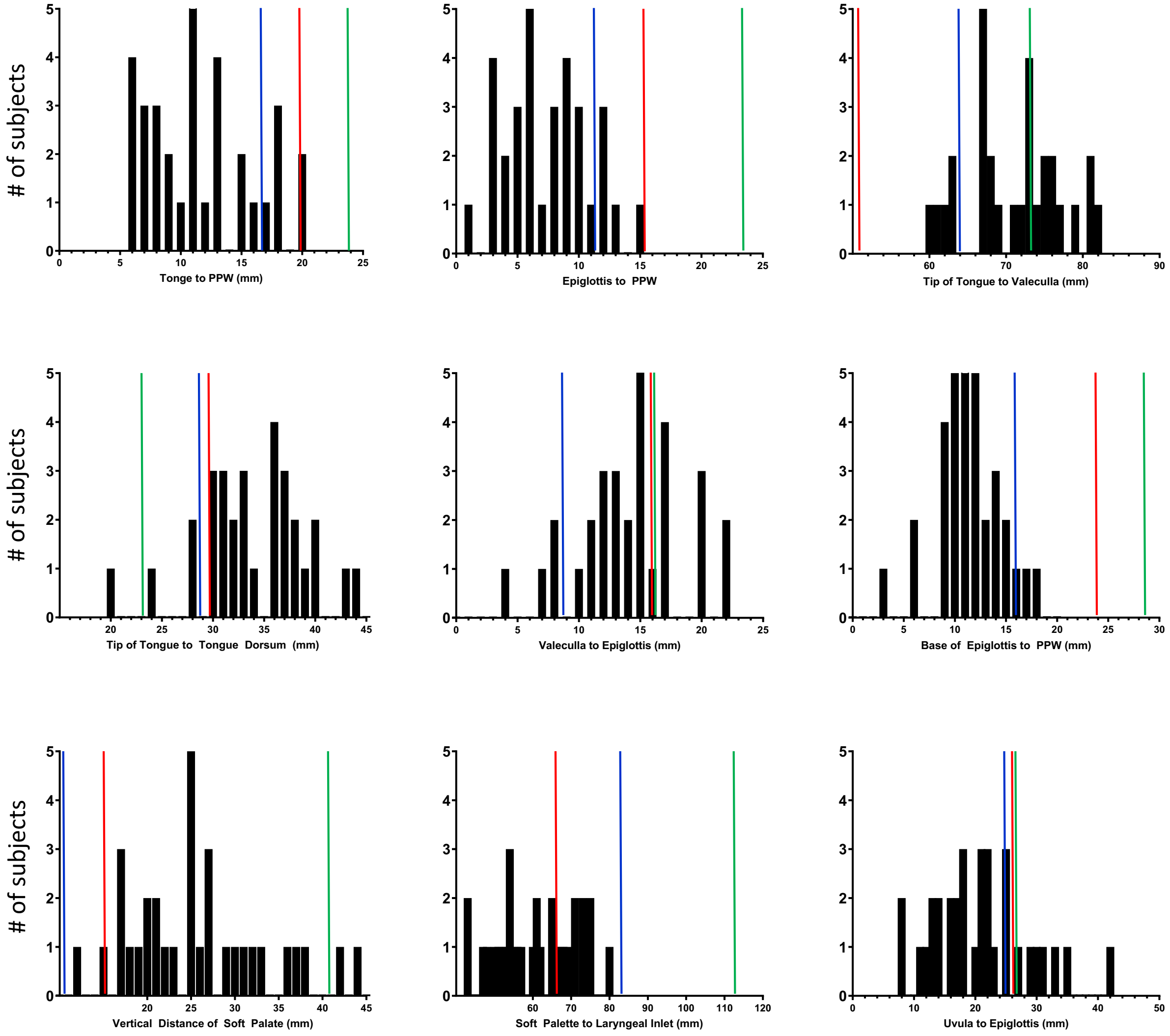

Figure 2 Distribution of participants' measurements with manikin measurements overlaid to view the disparity between subjects and manikins. Red, SynDaver; green, Laerdal; blue, AirSim. PPW, posterior pharyngeal wall.

Discussion

Our data demonstrate that low-fidelity airway manikins (SynDaver, Laerdal and AirSim) are not anatomically accurate when compared with adult human subjects. In particular, the airway space between the epiglottis and posterior pharyngeal wall, through which airway devices must pass, was too large in all three manikins. Contrary to Schalk et al. [15], who concluded that the Laerdal replicated human anatomy somewhat satisfactorily, in Figure 3 we show that Laerdal is a visual outlier compared with 33 human subjects when plotting the first two dimensions from multidimensional scaling. The current study adds valuable insight beyond previous studies in that: we focused specifically on low-fidelity manikins which, due to their lower cost, are most often used for device development and scientific research; we used two human populations, the first from the University of Michigan and the second from previously published literature to ensure diversity and robustness in our comparisons; we included measures that could not be captured by CT scan; and, this is the first study, to our knowledge, that included the SynDaver manikin, one of the most recently available.

Studies often use time to securement and first-pass success rate as primary endpoints, both of which may be positively impacted by a larger than average manikin airway size [23]. Additionally, studies often seek to determine provider device preference, which again may be influenced by improper airway anatomy [14]. Scutt et al. [24] found that manikin selection can greatly influence airway device performance, and specifically, slight alterations in manikin anatomy can impact results.

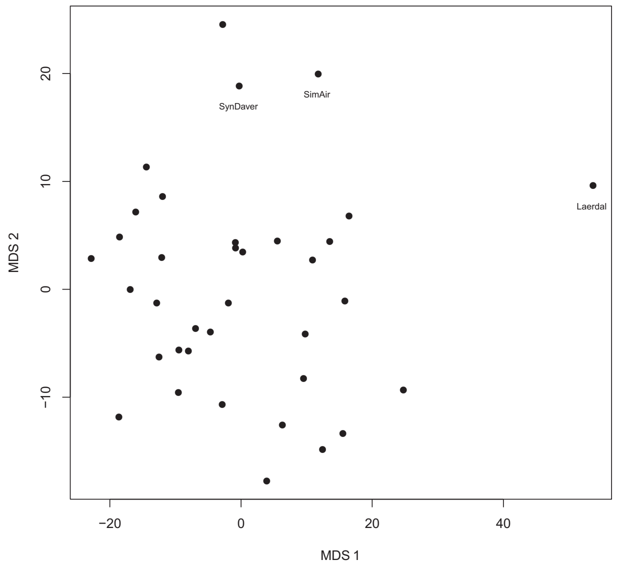

Figure 3 First and second dimensions from multidimensional scaling using Euclidean distances. Manikins (labelled) are clear outliers among participants.

Table 2 Human measurements from the literature compared with hand measurements for SynDaver, Laerdal and AirSim manikins. Manikin values are mean and percentile score to human measurement where 0.5 represents anatomic accuracy.

Measurement	Mean (SD) Published human values	Value (percentile) SynDaver	 Laerdal	 AirSim
Height mouth opens; cm [17]	4.78 (0.83)	5.98 (0.93)	4.52 (0.38)	8.10 (> 0.99)
First tracheal ring width; mm [18]	1.9 (0.6)	2.44 (0.82)	1.52 (0.26)	1.92 (0.51)
Trachea length; cm [19]	8.6 (1.1)	9.54 (0.80)	3.27 (< 0.01)	7.83 (0.24)
Neck circumference; cm [20]	36.6 (3.5)	33.4 (0.18)	43.55 (0.98)	41.23 (0.91)
Tongue to PPW; mm [16]	16.0 (0.7)	23.0 (> 0.99)	35.7 (> 0.99)	23.2 (> 0.99)
Epiglottis to PPW; mm [16]	9.0 (0.4)	16.7 (> 0.99)	35.4 (> 0.99)	14.8 (> 0.99)

PPW, posterior pharyngeal wall.

Interestingly, the distribution of human subjects as shown in Figure 2 was wide and variable for all measurements taken. In theory, this should have increased the likelihood that manikins would fall within the spread but, even with the wide distribution, they were still significantly outside the range. An additional study examining this variability in human subjects is currently underway. An understanding of the spatial relationship between external features and internal compartments of the airway could guide invasive procedures necessary for airway management. Analytic morphomics, a system developed by us [25], defines variability within a population so that location, dimensions and orientation of the targeted body compartments can be predicted from external features. This prediction would improve key airway objectives such as oral-tracheal intubation, surgical airway and central venous access. Using these type of data may allow for future development of autonomous or unmanned medical capabilities that require fundamental anatomic knowledge.

The study did not have the power to separate out men and women, but many manufacturers produce male and female specific trainers. A larger sample size of participants would allow for sub-setting by sex or other characteristics; the manikins may better represent specific sub-sets of humans. However, the current study samples both sexes and a range of ages and BMIs from the human population, all of which contribute to the variance of the sample. It is unlikely, given this increased variance, that the conclusion that measures from manikins with extreme measurements (> 99 or < 1 percentile) are not anatomically representative of humans, would change with increased sample size. Even several of the measurements that fall within two standard deviations from the mean could better represent the central tendency of the human population. For example, the measures for the uvula to the epiglottis fall within 1 SD of the mean for all three manikins, but all three also fall outside the 95%CI of the mean, which means that while the manikins fall within the normal range in humans, their measurements are not representative of the central tendency. The comparison of manikin measures with published airway values from different human sub-sets supports this as, for most measures, the manikins were not representative of the central tendency in the sample.

While the results of this study are vital to the eventual production of anatomically correct manikins, it has limitations. We focused on only three specific airway trainers. There are numerous airway manikins available and we chose ours based on common usage within our institution and those most often mentioned in scientific research studies. Other manikins may better represent human morphometries. Additionally, manikins are stiff and non-compliant rather than soft and compliant. In order for industry manufacturers to produce trainers that are more realistic, parameters such as elasticity and rigidity of airway structures need to be investigated thoroughly. Lastly, secretions, bleeding and gag reflexes are difficult to simulate given their wide range of presentations in difficult airways.

Acknowledgements

The authors would like to acknowledge the United States Army Medical Research and Development Command Tactical Combat Casualty Care Research Program. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Army Medical Department, Department of the Army, DoD or the US Government. No external funding or competing interests declared.

References

1. Cook TM, Woodall N, Frerk C. A national survey of the impact of NAP4 on airway management practice in United Kingdom hospitals: closing the safety gap in anaesthesia, intensive care and the emergency department. *British Journal of Anaesthesia* 2016; **117**: 182–90.
2. Je S, Cho Y, Choi HJ, Kang B, Lim T, Kang H. An application of the learning curve-cumulative summation test to evaluate training for endotracheal intubation in emergency medicine. *Emergency Medicine Journal* 2015; **32**: 291–4.
3. Sakles JC, Chiu S, Mosier J, Walker C, Stolz U. The importance of first pass success when performing orotracheal intubation in the emergency department. *Academic Emergency Medicine* 2013; **20**: 71–8.
4. Sebesta J. Special lessons learned from Iraq. *Surgical Clinics of North America* 2006; **86**: 711–26.
5. Sayre MR, Sakles JC, Mistler AF, Evans JL, Kramer AT, Pancioli AM. Field trial of endotracheal intubation by basic EMTs. *Annals of Emergency Medicine* 1998; **31**: 228–33.
6. Garza AG, Gratton MC, Coontz D, Noble E, Ma OJ. Effect of paramedic experience on orotracheal intubation success rates. *Journal of Emergency Medicine* 2003; **25**: 251–6.
7. Stewart RD, Paris PM, Pelton GH, Garretson D. Effect of varied training techniques on field endotracheal intubation success rates. *Annals of Emergency Medicine* 1984; **13**: 1032–6.
8. Sagarin MJ, Barton ED, Chng YM, Walls RM. Airway management by US and Canadian emergency medicine residents: a multicenter analysis of more than 6,000 endotracheal intubation attempts. *Annals of Emergency Medicine* 2005; **46**: 328–36.
9. Peters J, van Wageningen B, Hendriks I, et al. First-pass intubation success rate during rapid sequence induction of prehospital anaesthesia by physicians versus paramedics. *European Journal of Emergency Medicine* 2015; **22**: 391–4.
10. Cooper JB, Taqueti VR. A brief history of the development of mannequin simulators for clinical education and training. *Postgraduate Medical Journal* 2008; **84**: 563–70.
11. Mabry RL, Frankfurt A. Advanced airway management in combat casualties by medics at the point of injury: a sub-group analysis of the reach study. *Journal of Special Operations Medicine* 2011; **11**: 16–9.

12. Jackson KM, Cook TM. Evaluation of four airway training manikins as patient simulators for the insertion of eight types of supraglottic airway devices. *Anaesthesia* 2007; **62**: 388–93.
13. March JA, Tassey TE, Resurreccion NB, Portela RC, Taylor SE. Comparison of the I-gel supraglottic and king laryngotracheal airways in a simulated tactical environment. *Prehospital Emerggency Care* 2018; **22**: 385–9.
14. Schauer SG, Fernandez JRD, Roper JL, et al. A randomized cross-over study comparing surgical cricothyrotomy techniques by combat medics using a synthetic cadaver model. *American Journal of Emergency Medicine* 2018; **36**: 651–6.
15. Schalk R, Eichler K, Bergold MN, et al. A radiographic comparison of human airway anatomy and airway manikins – implications for manikin-based testing of artificial airways. *Resuscitation* 2015; **92**: 129–36.
16. Schebesta K, Hupfl M, Rossler B, Ringl H, Muller MP, Kimberger O. Degrees of reality: airway anatomy of high-fidelity human patient simulators and airway trainers. *Anesthesiology* 2012; **116**: 1204–9.
17. Khare N, Patil SB, Kale SM, Sumeet J, Sonali I, Sumeet B. Normal mouth opening in an adult Indian population. *Journal of Maxillofacial and Oral Surgery* 2012; **11**: 309–13.
18. Otoch JP, Minamoto H, Perini M, Carneiro FO, de Almeida Artifon EL. Is there a correlation between right bronchus length and diameter with age? *Journal of Thoracic Disease* 2013; **5**: 306–9.
19. Cinar U, Halezeroglu S, Okur E, Akif Inanici M, Kayaoglu S. Tracheal length in adult human: the results of 100 autopsies. *International Journal of Morphology* 2016; **34**: 232–6.
20. Schwab RJ, Gupta KB, Gefter WB, Metzger LJ, Hoffman EA, Pack AI. Upper airway and soft tissue anatomy in normal subjects and patients with sleep-disordered breathing. Significance of the lateral pharyngeal walls. *American Journal of Respiratory and Critical Care Medicine* 1995; **152**: 1673–89.
21. Gower JC. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 1966; **53**: 325–38.
22. Mardia KV. Some properties of classical multi-dimensional scaling. *Communications in Statistics – Theory and Methods* 1978; **7**: 1233–41.
23. Rai MR, Popat MT. Evaluation of airway equipment: man or manikin? *Anaesthesia* 2011; **66**: 1–3.
24. Scutt S, Clark N, Cook TM, et al. Evaluation of a single-use intubating videoscope (Ambu aScope) in three airway training manikins for oral intubation, nasal intubation and intubation via three supraglottic airway devices. *Anaesthesia* 2011; **66**: 293–9.
25. Wang C, Vainshtein JM, Veksler M, et al. Investigating the clinical significance of body composition changes in patients undergoing chemoradiation for oropharyngeal cancer using analytic morphomics. *Springerplus* 2016; **5**: 429.