This is an Accepted Manuscript for *The Journal of Laryngology & Otology*
DOI: 10.1017/S0022215124002226

Face and content validation of additive manufacturing temporal bone specimens

Jesslyn Clarance Lamtara, B.BIOMED(Hons)

Department of Surgery (Otolaryngology), University of Melbourne, Victoria, Australia.

Sudanthi Wijewickrema, PhD

Department of Surgery (Otolaryngology), University of Melbourne, Victoria, Australia.

Jean-Marc Gerard, Dr, A/Prof

Department of Surgery (Otolaryngology), University of Melbourne, Victoria, Australia.

The Royal Eye and Ear Hospital, Melbourne, Victoria, Australia

Stephen O'Leary, Prof

Department of Surgery (Otolaryngology), University of Melbourne, Victoria, Australia.

The Royal Eye and Ear Hospital, Melbourne, Victoria, Australia

Corresponding author:

Jesslyn Clarance Lamtara

Email: jlamtara@student.unimelb.edu.au

Address: Department of Surgery (Otolaryngology)

University of Melbourne

Level 5, The Royal Eye and Ear Hospital

32 Gisborne Street, East Melbourne, Victoria 3002

Australia

Phone: +61 3 9929 8365

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Abstract

Introduction. Otology training solely using cadavers is challenging due to scarcity and high costs. The use of additive manufacturing (AM) technology is a promising alternative. This study aimed at qualitatively validating new AM temporal bone specimens as to their realism and ability to train surgical skills.

Material and methods. Three AM models generated using cadaveric temporal bones were evaluated. Three otologists with experience as trainers dissected and evaluated each specimen.

Results. The AM specimens scored an average of 4.26 ± 0.72 (out of 5) points and received positive feedback. The agreement between the 3 expert raters was high (intra-class correlation coefficient of 0.745).

Conclusion. Results suggested that the AM temporal bones were able to faithfully reproduce a training experience similar to that on cadaveric temporal bones. Further studies that investigate the effectiveness of these specimens in training surgical skills are needed before integrating them into surgical training curricula.

Key words: Additive manufacturing, 3D printing, temporal bone, mastoidectomy, otology, training, face validity, content validity, medical education.

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Introduction

Temporal bone surgery is complex, as it requires navigation around critical structures such as dura of the middle and posterior fossa, sigmoid sinus, carotid artery, facial nerve, and the labyrinthine[1]. The gold standard of temporal bone surgery, as in any other surgical specialty, is cadaveric dissection[2]. Cadaveric dissection provides a profound visual and tactile experience that mimics surgery on a living human body[2]. However, the sole use of this method of surgical training has become increasingly difficult due to issues such as scarcity of cadavers, high costs, limited availability of equipped dissection laboratories, and possibility of exposure to biological risks[3, 4]. Furthermore, exposure to paediatric and rare pathologies, as well as a range of anatomical variation, is not always possible with cadaveric dissection.

Therefore, supplementary methods of surgical training have been actively investigated to ensure trainees have achieved a sufficient level of skills before handling scarce resources. Leading among these supplementary methods are: computer-based simulation such as virtual reality (VR) and the use of additive manufacturing (AM), or 3D printed, specimens. While VR offers a repeatable, low-cost experience, it is not able to sufficiently simulate the surgical experience of a real dissection. On the other hand, AM specimens provide a more realistic dissection experience along with repeatability. However, the long-term cost of AM surgical training is higher as the specimens are not reusable. Both forms of simulation offer the ability to create standardized surgical curricula, with exposure to different anatomies, rare pathologies, and paediatric specimens.

Various studies spanning over 2 decades have established the face and content validity of VR simulation and its effectiveness in improving surgical skills in temporal bone surgery[5-11]. However, as the use of AM models is a more recent advance, not as many studies exist that investigate its effectiveness in this context.

Previous work has investigated the creation of AM models that realistically mimic the appearance and tactility of a cadaveric temporal bone[3, 12-16]. For example, Mowry et. al[17] and Frithioff et. al[18] conducted reviews of 3D printed temporal bone models, comparing their software applications, 3D printers, and material. They found that stereolithography (SLA) printers with powder and resin as the printing materials provided best results. Validity of AM temporal bone models were typically performed using face and content validity questionnaires[3, 12-16]. These qualitative evaluations largely used the opinion of surgical registrars with a few employing expert surgeons[18]. In a recent systematic review, Frithioff et.al[18] concluded that, although most studies reported positive attitudes toward the models and their potential for training, the educational quality of such validations was low (i.e., Kirkpatrick level 1)[19]. Therefore, stronger evidence is required as to the effectiveness of AM in the training of temporal bone surgery.

However, qualitative validation of AM specimens cannot be dismissed out of hand, as it paves the way to investigations on their training efficacy and subsequent integration into temporal bone surgery curricula. Also, as additive manufacturing technology evolves rapidly, it is important that the AM models are continuously validated as to their ability to faithfully reproduce the anatomical structures of the temporal bone and overall drilling experience[18]. As such, in this study, we investigated the face and content validity of new AM temporal bone specimens developed and manufactured by Fusetec

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

(Adelaide, South Australia). In contrast to previous work that only investigated healthy normal temporal bone specimens, we validated 3 different cases: healthy normal, healthy highly pneumatised, and healthy sclerotic bones. To obtain more reliable feedback, and minimise individual bias, we employed 3 senior consulting otologists who are involved in surgical training for the validation.

Materials and Methods

AM temporal bone models

The AM temporal bone models obtained from Fusetec were advanced manufactured based on high resolution axial computed tomography (CT) scans of cadaveric temporal bones. They were segmented on Materialise Mimics software, then converted to STL files in Netfabb for mesh cleaning and customising specific engineering features. Materials were selected under surgical guidance from previously document dog bones tensile tests. The bone and anatomical structures were produced using multiple materials using 0.0125mm slices with a proprietary voxel-based software integration. All soft tissue and bony structures of the temporal bone were represented: inner and outer auditory canal, labyrinthine, tympanic membrane, ossicles, facial nerve, chorda tympani, sigmoid sinus, carotid artery and dura mater. To facilitate realistic haptic feedback on bone and soft tissue, the models were constructed with Shore Hardness of 83-86 and 28-23 respectively. The 3 cases of temporal bone models used in the study were produced with different air cell composition to replicate normal, highly pneumatised, and sclerotic bones. Figure 1 shows axial cross-sections of the 3 temporal bone models used.

Study design

Ethics approval was obtained from the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital (HREC number 24/1599HL). All participants provided written consent.

Three senior otologists who are involved in teaching temporal bone surgery were recruited. The surgeons' experience in training varied: 6, 5, and 5 years for raters 1, 2, and 3 respectively. Each surgeon performed temporal bone dissections to evaluate the face and content validities of the 3 cases of the AM temporal bone specimens. The dissections were conducted at the Victorian Eye and Ear Hospital Temporal Bone Laboratory, using standard operating theatre equipment (microscope, micro-drills, and irrigation-suction systems) and all necessary personal protective equipment (gowns, gloves, masks, and eye protection).

After performing each dissection, each surgeon completed a questionnaire assessing their experience with the specimen. The questionnaire developed by Da Cruz and Francis[12] was used for this purpose, selected by an expert otologist due to its ability to assess anatomical / drilling realism and elements addressing temporal bone dissection. This questionnaire consisted of 23 questions in 4 categories (anatomical realism, usefulness as a training tool, task-based usefulness, and overall reactions), based on a 5-point Likert scale (1-strongly disagree to 5-strongly agree). In addition, the surgeons were also asked to provide any feedback/comments outside the questionnaire.

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

The questionnaire responses were tabulated in Microsoft Excel and average scores and standard deviations for all specimens, as well as those for each specimen, were calculated. Intraclass correlation coefficient (ICC) between the 3 raters was calculated using SPSS version 29 (IBM, Chicago, IL) to test for inter-rater reliability. In addition, the questionnaire scores were also compared to the scores of existing 3D printed temporal bone specimens. To enable comparison across different validation scales, 4 categories of evaluation used in the literature for AM validation was identified: anatomical realism, drilling realism, basic surgical skills, mastoid surgery skills, and skill transfer to real patients. For each study, the items belonging to each category were identified from the scoring scale used and the average scores were calculated based on the number of items and number of participants. The scores were then normalised so that each category represented a score out of 100.

Results

Overall Results

Figure 2 illustrates the 3 drilled AM specimens. Table 1 shows the responses for all 3 specimens from the 3 raters. The ICC of the overall specimens for 3 raters was 0.745. The Bland-Altman plots in Figure 3 illustrate the agreement between each pair of raters. Bland-Altman plots, also known as difference plots, are a convenient way to assess the agreement between two sets of measurements[20]. The y-axis shows the difference between the two paired measurements and the x-axis represents the average of these measures. Mean and 95% confidence intervals of the differences are also plotted (upper and lower limits in Figure 3). An ideal agreement is zero difference[21]. Each

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

point/bubble represents an instance of the difference in a pair of ratings against their average.

Results for the 3 individual specimens

The score of the healthy normal bone specimen from 3 raters (mean ± standard deviation) was 4.20 ± 0.77, while that of healthy highly pneumatised bone specimen was 4.33 ± 0.66. The healthy sclerotic bone specimen received a score of 4.23 ± 0.73. ICC of the 3 raters for the healthy normal, highly pneumatised, and sclerotic bone specimens were 0.820, 0.652, and 0.747, respectively. Figures 4, 5 and 6 illustrate the results of the face and content validity questionnaire from 3 raters for the above specimens respectively.

Overall, raters commented that the bone specimens had a very realistic drilling experience, including drilling pitch and residue. The raters' opinion was that the colour of the specimens could be whiter to allow the structures to be pinker rather than yellow. The healthy normal bone specimen was identified as a good specimen for advanced cases because the cortical mastoidectomy was very easy, but the facial recess was very tight. Meanwhile, raters commented that the healthy sclerotic bone specimen was a good representation for difficult anatomy, specifically the tegmen and sigmoid sinus. This bone specimen also produced a higher drilling pitch compared to the other 2 specimens.

Comparisons with other existing 3D temporal bone models

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Table 2 compares the questionnaire scores of this study and others that used the same validation scale[12]. Table 3 shows a category-wise comparison of existing AM temporal bone evaluations across different validation scales, normalised as discussed above.

Discussion

Cadaveric dissection remains the gold standard in surgical training, including in otology, despite issues, such as disease transmission, maintenance cost, and limited availability[3]. Alternatives, such as VR simulation and AM models have been actively investigated to address the drawbacks of the sole use of cadaveric dissection in training temporal bone surgery.

This study evaluated the face and content validity of 3 different cases of 3D printed temporal bone models manufactured by Fusetec (Adelaide, South Australia) using the questionnaire developed by Da Cruz and Francis [12]. We used senior ENT surgeons experienced in training registrars for the validation of the specimens rather than ENT trainees used in some previous studies. This ensured that the results were of a higher quality as their knowledge of training requirements and what is required of a training specimen is more reliable.

Overall performance

These models received highly positive responses from the 3 raters, with an average of 4.26 ± 0.72 (out of 5) indicating that experts were of the opinion that they would offer a similar surgical training experience as cadaveric dissection. These models rated highest for their usefulness for teaching cortical mastoidectomy (4.89 out of 5). On the other hand, item colour contrast was rated the lowest (3.22 out of 5). The raters suggested the bone should be whiter to allow the structures to be pinker rather than yellow. Raters also highlighted that the models produced very realistic residue and drilling pitch.

Case-wise performance

There were 3 different cases used in this study, none with any pathologies (normal, pneumatised, and sclerotic) temporal bone specimens. Each questionnaire item of each specimen received high responses from the 3 raters (3 and above), with scores of 4.20 ± 0.72, 4.33 ± 0.66, and 4.23 ± 0.73 for the 3 specimens respectively. In addition, the agreement levels of the 3 raters for each specimen were also high, Raters commented that the high drilling pitch on the sclerotic temporal bone specimen was quite realistic.

Inter-rater reliability

The agreement level among 3 raters was high (overall ICC = 0.745), with the lowest agreement being for the highly pneumatised specimen (ICC = 0.652) and the normal and sclerotic specimens receiving ICC scores of 0.820 and 0.747 respectively. Additionally, according to the Bland-Altman plots, rater 1 was observed to be consistently most lenient and rater 3 being the strictest in their rating. As the 3 raters were of similar experience levels, this difference in scoring is likely due to individual differences and standards.

Comparison with existing literature

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

When comparing the face and content validity results of Da Cruz and Francis [12], Chien et.al [14] and this study (Table 2), which utilised the same questionnaire, this study achieved the highest average score, followed closely by Da Cruz and Francis [12] and lastly, Chien et.al (2021) [14]. From Table 3, which compared the face and content validity of existing AM temporal bone specimens using various questionnaires, it can be seen that the scores for anatomical realism decreased from 2015 to 2024, while that of the other categories either increased or remained the same. This seems to contradict the fact that with the advancement of additive manufacturing technology, better AM models with higher resolution are being produced. This decrease could be because of higher expectations on the part of the participants, as they are increasingly exposed to improved AM specimens through the years.

Comparison of recent AM models and printing technologies

The recent studies conducted by Mowry et.al [17] and Iannella et.al [3] are the most comparable to the current study, as they used more advanced technology as well as the more reliable assessments of experts. According to Mowry et.al [17], the highest scoring models were produced with FormLabs Form 2 and Zcorp 650 stereolithography (SLA) printers, which offer a range of resin material but are limited to printing one material at a time. Likewise, Iannella et.al [3] used a Photon mono x 4k SLA printer to create their TB models. Additionally, these machines are designed for engineering hard plastic prototypes, and some require extensive post-printing treatments. On the other hand, Fusetec AM technology allows printing of multiple materials and adjustment of shore hardness to 70A. Different materials can be selected and allocated in almost any proportion digitally prior to printing. As such, a large range of shore values, texture,

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

colour and density can be produced that simulate characteristics of different tissue types. Moreover, post-printing treatment is only required to wash out support material. As such, Fusetec AM models offer a high level of flexibility in defining colours, textures, and haptic properties.

Limitations and future work

The comparisons in Table 3 have been divided into categories, and normalised, so that the scores of different studies that used different validation questionnaires could be compared. This, in addition to the differences in rater experience, as well as the rater numbers, add bias to the comparison results.

As observed in Frithioff et. al [18], qualitative studies such as the one discussed here, are not sufficient by themselves to validate the educational quality of AM models. Nevertheless, they are still important as they pave the way for higher quality validations. Indeed, the next step is to investigate the effectiveness of these AM models in training temporal bone surgery skills.

Once this has been established, a simulation-based curriculum that incorporates VR and AM training should be designed and validated, in order to take advantage of the benefits offered by both technologies.

Summary

- Sole use of cadavers in temporal bone surgery training is impractical because of resources scarcity and high costs.
- The proposed additive manufacturing (AM) temporal bone specimens were able to faithfully reproduce a training experience similar to that on cadaveric temporal bone.
- Further studies regarding the effectiveness of these specimens in training surgical skills are needed.

Conclusion

This study successfully established the face and content validity of 3 different AM temporal bone specimens manufactured by Fusetec. Further studies regarding the effectiveness of these models in improving trainees' temporal bone surgical skills are needed so that these specimens could be integrated into surgical curricula.

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

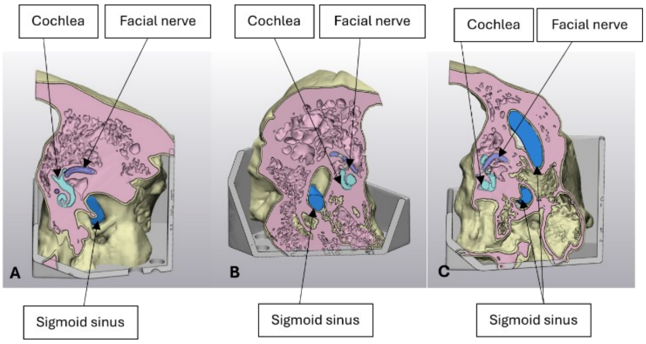

Figure 1. Axial cross-sections through the cochlea of the 3 cases of AM temporal bone: healthy normal bone (A), healthy highly pneumatised bone (B) and healthy sclerotic bone (C). The cutting plane is in pink, while the structures cochlea, facial nerve, and sigmoid sinus are coloured in light blue, purple, and dark blue respectively. Note the difference in pneumatisation of the 3 specimens.

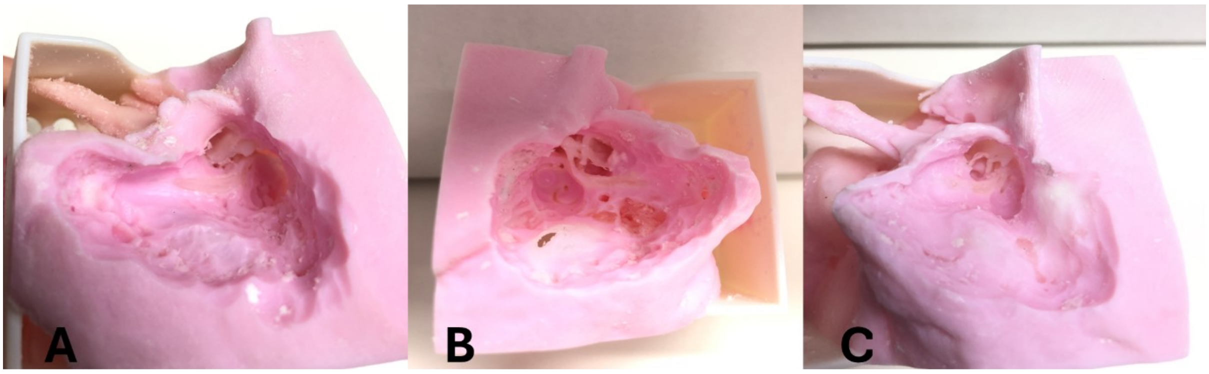

Figure 2. The 3 cases of AM temporal bones after drilling: healthy normal bone (A), healthy highly pneumatised bone (B) and healthy sclerotic bone (C).

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

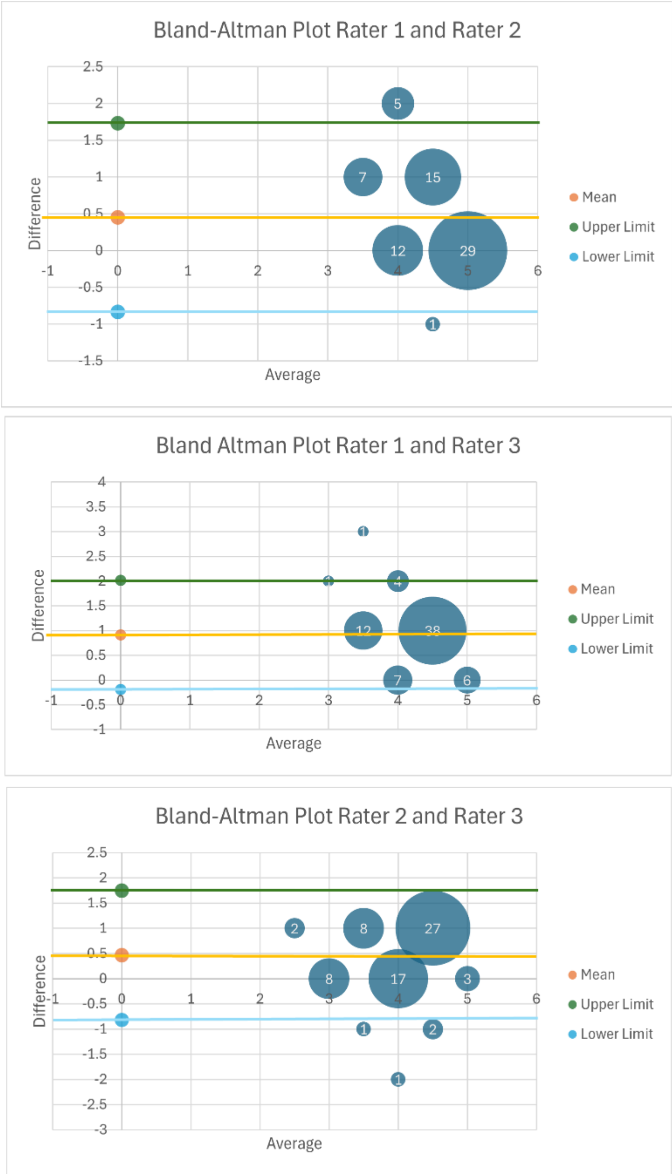

Figure 3. Bland-Altman plots of the overall specimens result among 3 raters. Numbers in the bubble represent the number of repetitions of agreement between each pair of raters at that point (the number of instances a given difference in the ratings with respect to their mean occurred).

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

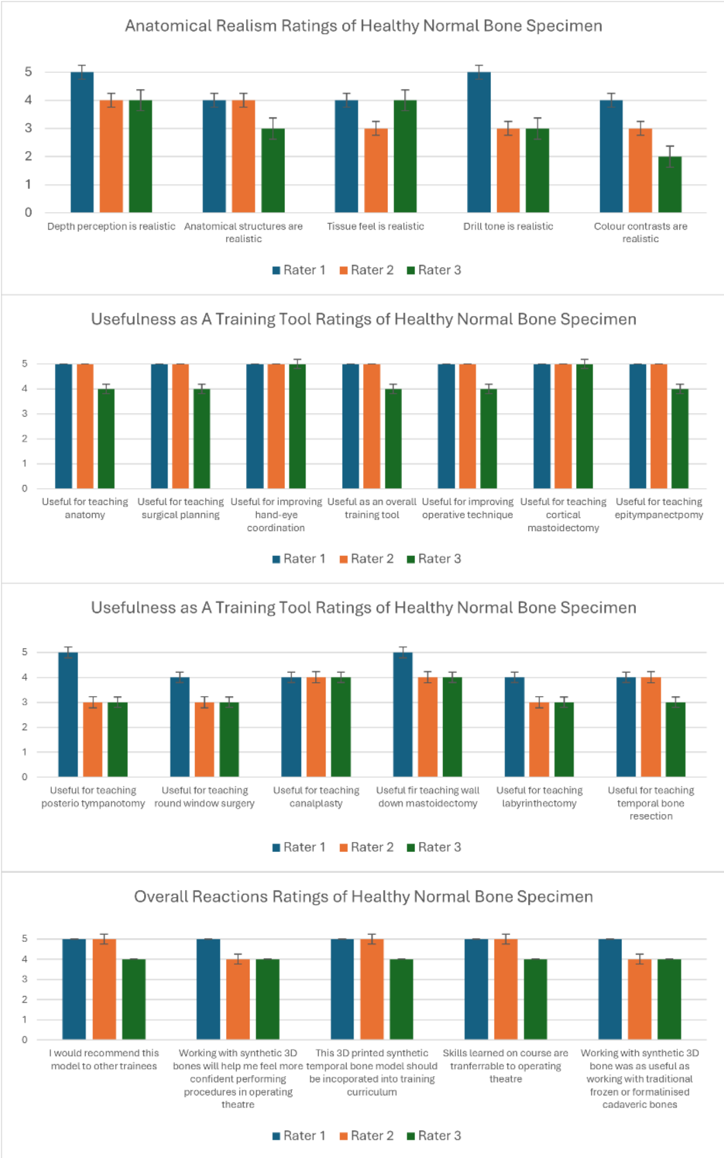

Figure 4. Face and content validity questionnaire responses of healthy normal bone specimen from 3 raters.

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

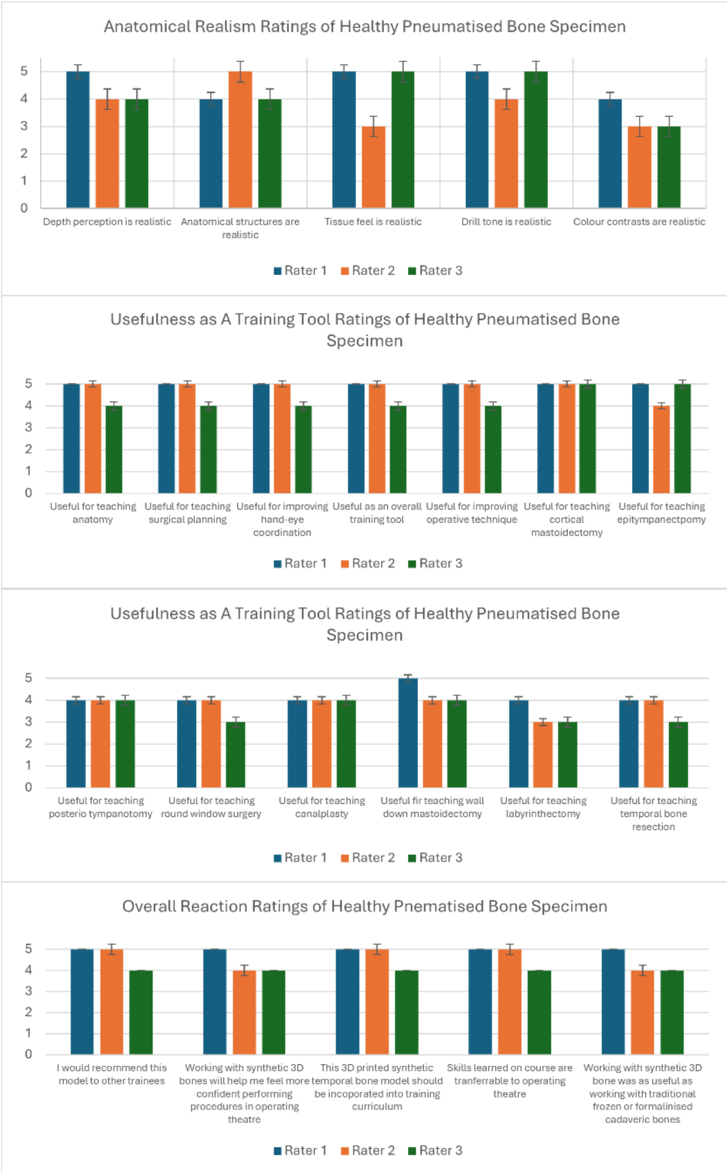

Figure 5. Face and content validity questionnaire responses of healthy highly pneumatised bone specimen from 3 raters.

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Figure 6. Face and content validity questionnaire responses of healthy sclerotic bone specimen from 3 raters.

Acknowledgements

The authors thank Fusetec (Adelaide, South Australia) for providing the AM temporal bone specimens for this study.

Fundings: None

Competing Interests: The authors declare none.

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

References

1 Zagzoog N, Yang VXD. State of Robotic Mastoidectomy: Literature Review. 2018. p. 347-51

2 Hu M, Wattchow D, de Fontgalland D. From ancient to avant-garde: a review of traditional and modern multimodal approaches to surgical anatomy education. *ANZ Journal of Surgery* 2018;**88:**146-51

3 Iannella G, Pace A, Mucchino A, Greco A, De Virgilio A, Lechien JR, et al. A new 3D-printed temporal bone: ‘the SAPIENS’—specific anatomical printed-3D-model in education and new surgical simulations. *European Archives of Oto-Rhino-Laryngology: and Head & Neck* 2024**:**1-10

4 Frithioff A, Sørensen MS, Andersen SAW. European status on temporal bone training: a questionnaire study. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 2018;**275:**357-63

5 O'Leary SJ, Tykocinski M, Dahm M, Pyman B, Hutchins MA, Stevenson DR, et al. Validation of a networked virtual reality simulation of temporal bone surgery. *Laryngoscope* 2008;**118:**1040-6

6 Al-Noury K. Virtual Reality Simulation in Ear Microsurgery: A Pilot Study. *Indian Journal of Otolaryngology Head and Neck Surgery* 2012;**64:**162-6

7 Andersen SA, Foghsgaard S, Konge L, Caye-Thomasen P, Sorensen MS. The effect of self-directed virtual reality simulation on dissection training performance in mastoidectomy. *Laryngoscope* 2016;**126:**1883-8

8 Andersen SAW, Cayé-Thomasen P, Sørensen MS. Mastoidectomy performance assessment of virtual simulation training using final-product analysis. *The Laryngoscope* 2015;**125:**431-5

9 Sorensen MS, Mosegaard J, Trier P. The visible ear simulator: a public PC application for GPU-accelerated haptic 3D simulation of ear surgery based on the visible ear data. Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology. United States: Lippincott Williams & Wilkins; 2009. p. 484-7

10 Varoquier M, Hoffmann CP, Perrenot C, Tran N, Parietti-Winkler C. Construct, Face, and Content Validation on Voxel-Man® Simulator for Otologic Surgical Training. *International Journal of Otolaryngology* 2017**:**1-8

11 Copson B, Wijewickrema S, Zhou Y, Piromchai P, Briggs R, Bailey J, et al. Supporting skill acquisition in cochlear implant surgery through virtual reality simulation. Cochlear implants international. England: Taylor & Francis; 2017. p. 89-96

12 Da Cruz MJ, Francis HW. Face and content validation of a novel three-dimensional printed temporal bone for surgical skills development. *The Journal of Laryngology and Otology* 2015;**129:**S23-S9

13 Wong V, Unger B, Pisa J, Gousseau M, Westerberg B, Hochman JB. Construct Validation of a Printed Bone Substitute in Otologic Education. *Otology and Neurotology* 2019;**40:**E698-E703

14 Chien WW, da Cruz MJ, Francis HW. Validation of a 3D-printed human temporal bone model for otology surgical skill training. *World Journal of Otorhinolaryngology - Head and Neck Surgery* 2021;**7:**88-93

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

15 Frithioff A, Frendø M, Foghsgaard S, Mikkelsen PT, Sørensen MS, Andersen SAW, et al. 3-D-Printed Models for Temporal Bone Training: A Validity Study. *Otology and Neurotology* 2023;**44:**E497-E503

16 Gadaleta DJ, Huang D, Rankin N, Hsue V, Sakkal M, Bovenzi C, et al. 3D printed temporal bone as a tool for otologic surgery simulation. *American Journal of Otolaryngology* 2020;**41:**102273

17 Mowry SE, Jabbour N, Rose AS, Wiet GJ, Svrakic M, Zopf DA, et al. Multi-institutional Comparison of Temporal Bone Models: A Collaboration of the AAO-HNSF 3D-Printed Temporal Bone Working Group. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 2021;**164:**1077-84

18 Frithioff A, Frendo M, Pedersen DB, Sorensen MS, Wuyts Andersen SA. 3D-Printed Models for Temporal Bone Surgical Training: A Systematic Review. *OTOLARYNGOLOGY-HEAD AND NECK SURGERY* 2021;**165:**617-25

19 Cook DA, Reed DA. Appraising the Quality of Medical Education Research Methods: The Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education. *Academic Medicine* 2015;**90:**1067-76

20 Giavarina D. Understanding Bland Altman analysis. *Biochemia Medica* 2015;**25:**141-51

21 Doğan NÖ. Bland-Altman analysis: A paradigm to understand correlation and agreement. *Turkish Journal of Emergency Medicine* 2018;**18:**139-41

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Table 1. Face and content validity questionnaire responses of overall specimens from the 3 raters.

Domain	**Subdomain**	**Average score**	**standard deviation**
Anatomical realism	Depth perception is realistic	4.33	0.50
	Anatomical structures are realistic	4.00	0.71
	Tissue feel is realistic	3.89	0.93
	Drill tone is realistic	3.89	1.17
	Colour contrasts are realistic	3.22	0.67
Usefulness as a training tool	Useful for teaching anatomy	4.67	0.50
	Useful for teaching surgical planning	4.67	0.50
	Useful for improving hand-eye coordination	4.78	0.44
	Useful as an overall training tool	4.67	0.50
	Useful for improving operative technique	4.67	0.50
Task-based usefulness	Useful for teaching cortical mastoidectomy	4.89	0.33
	Useful for teaching epitympanectomy	4.56	0.53
	Useful for teaching posterior tympanotomy	3.78	0.67
	Useful for teaching round window surgery	3.67	0.50
	Useful for teaching canalplasty	4.00	0

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

	Useful for teaching wall down mastoidectomy	4.33	0.50
	Useful for teaching labyrinthectomy	3.44	0.53
	Useful for teaching temporal bone resection	3.67	0.50
Overall reactions	I would recommend this model to other trainees	4.67	0.50
	Working with synthetic 3D bones will help me feel more confident performing procedures in operating theatre	4.33	0.50
	This 3D printed synthetic temporal bone model should be incorporated into training curriculum	4.67	0.50
	Skills learned on course are transferable to operating theatre	4.67	0.50
	Working with synthetic 3D bone was as useful as working with traditional frozen or formalinised cadaveric bones	4.44	0.53

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Table 2. Comparison of face and content validity questionnaire scores of Da Cruz and Francis[12], Chien et.al[14] and this study which used the same questionnaire.

Domain	**Subdomain**	**Da Cruz and Francis (2015)[12,a]**	**Chien et.al (2021)[14,a]**	**This study (2024)[a]**
Anatomical realism	Depth perception is realistic	4.7	4	4.33
	Anatomical structures are realistic	4.3	3.53	4.00
	Tissue feel is realistic	4	3.71	3.89
	Drill tone is realistic	3.7	2.82	3.89
	Colour contrasts are realistic	3.2	2.47	3.22
Usefulness as a training tool	Useful for teaching anatomy	4.8	4.59	4.67
	Useful for teaching surgical planning	4.8	4.53	4.67
	Useful for improving hand-eye coordination	4.8	4.59	4.78
	Useful as an overall training tool	4.7	4.65	4.67
	Useful for improving operative technique	4.6	4.35	4.67

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

Task-based usefulness	Useful for teaching cortical mastoidectomy	4.8	4.59	4.89
	Useful for teaching epitympanectomy	4.2	3.87	4.56
	Useful for teaching posterior tympanotomy	4.2	3.73	3.78
	Useful for teaching round window surgery	3.7	3.47	3.67
	Useful for teaching canalplasty	3.5	3.33	4.00
	Useful for teaching wall down mastoidectomy	3.5	4.13	4.33
	Useful for teaching labyrinthectomy	4.3	3.87	3.44
	Useful for teaching temporal bone resection	3.7	3.33	3.67
Overall reactions	I would recommend this model to other trainees	4.8	4.25	4.67
	Working with synthetic 3D bones will help me feel more confident performing procedures in operating theatre	4.7	4.19	4.33
	This 3D printed synthetic temporal bone model should be	4.7	4.19	4.67

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

	incorporated into training curriculum			
	Skills learned on course are transferable to operating theatre	4.3	4.19	467
	Working with synthetic 3D bone was as useful as working with traditional frozen or formalinised cadaveric bones	3.5	3.19	4.44
AVERAGE		4.24	3.89	4.57

[a]Scores are out of 5

Table 3. *Comparison of face and validity scores between existing AM temporal bone specimens that used various questionnaires. The scores are all averages and have been adjusted to be out of 100%.*

	Da Cruz and Francis (2015) [12]	**Wong et.al (2019)** [13]	**Gadaleta et.al (2020)** [16]	**Chien et.al (2021)** [14]	**Mowry et.al (2021)** [17]	**Iannella et.al (2024)** [3]	**This study (2024)**
Participants	ENT residents	ENT residents	ENT residents	ENT residents	Senior ENT surgeons	ENT surgeons +	Senior ENT surgeons

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press

						neurosu rgeons	
Number of participants	9	19	10	17	4	5	3
Anatomical realism	83.5	74.57	66.5	70.3	69.78	72.5	77.22
Drilling realism	64	N/A	78.5	49.4	85.37	87.5	77.78
Basic surgical skills	94.8	N/A	N/A	90.84	53.4	92	93.78
Mastoid surgery skills	79.75	80.86	N/A	75.8	100	100	80.83
Skill transfer to real patients	86	82.81	87	80.04	N/A	N/A	91.11

Abbreviation N/A = Not Available

https://doi.org/10.1017/S0022215124002226 Published online by Cambridge University Press