

A training tool for clinicians in segmenting medical images to make 3D models

Soudeh Chegini, BMBCh, BDS, MA, Med, FRCS(OMFS),*,† Arpan Tahim,‡ Mingjun Liu,§ Eddie Edwards,† Yean Chooi,|| Matthew Clarkson,† and Clare Schilling¶

Introduction: 3D models produced from medical imaging can be used to plan treatment, design prosthesis, teach, and for communication. Despite the clinical benefit, few clinicians have experience of how 3D models are produced. This is the first study evaluating a training tool to teach clinicians to produce 3D models and reporting the perceived impact on their clinical practice.

Method: Following ethical approval, 10 clinicians completed a bespoke training tool, comprising written and video material alongside online support. Each clinician and 2 technicians (included as control) were sent 3 CT scans and asked to produce 6 fibula 3D models using open-source software (3Dslicer). The produced models were compared to those produced by the technicians using Hausdorff distance calculation. Thematic analysis was used to study the postintervention questionnaire.

Results: The mean Hausdorff distance between the final model produced by the clinicians and technicians was 0.65 mm \pm SD 0.54 mm. The first model made by clinicians took a mean time of 1 hour 25 minutes and the final model took 16:04 minutes (5:00–46:00 minutes). 100% of learners reported finding the training tool useful and will employ it in future practice.

Discussion: The training tool described in this article is able to successfully train clinicians to produce fibula models from CT scans. Learners were able to produce comparable models to technicians within an acceptable timeframe. This does not replace technicians. However, the learners perceived this training will allow them to use this technology in more cases, with appropriate case selection and they appreciate the limits of this technology.

Keywords: [MeSH]: medical education, three-dimensional imaging, Radiology Information Systems

INTRODUCTION

Medical imaging is widely used to plan procedures that are customized to the patients individual anatomy. Medical images can be segmented using software to isolate anatomy of interest and produce 3D models. 3D models can be useful in planning

complex procedures or in the decision-making of suitable procedures for each patient.¹ They are also used to examine the fit of custom implants. They have been shown to improve procedure accuracy and shorten procedure time. They can also be used as a communication tool with patients to allow informed consent and reduce anxiety.² Similarly, 3D models have been used to train clinicians.³ As the public's awareness of this technology increases, this may become the standard of care expected by patients.

Despite the benefits of 3D models in clinical practice, the use of 3D modeling is not yet standard practice.² In most cases the making of 3D models is outsourced to specialist technicians with few clinicians experienced in producing 3D models themselves. It is important for clinicians to understand the limitations of accuracy in how 3D models are produced to deploy these appropriately for clinical decision-making. Understanding a technique demands knowledge of the entirety of the process,⁴ and in turn, training clinicians in the use of new technology is known to promote its adoption in the wider profession.⁵

The availability of 3D printers in hospitals makes image segmentation a technology which is ripe for the surgeon to adopt. 3D models can also be appreciated on computers, smart devices and augmented and virtual reality headsets such as Oculus and Hololens.

Because of the advent of home computer systems, everyone has used software teaching materials. This is a well-studied field that was studied and deliberately applied to develop the tool in this study.⁶ Dreyfus-Dreyfus described a 5-stage process of skill acquisition, from novice to expert, which has been widely applied in medical education.⁷ This was applied in the evaluation of the training tool. Learning how to use a piece of software is a performed ability with results that can be quantitatively evaluated. Therefore, a model used for learning a skill was deemed the most appropriate.

The aim of this study is to evaluate the training tool designed for clinicians in the process of segmentation to produce 3D models. The questions are (1) the impact of the training tool on learning how to make 3D models and (2) the perceived value of

From the *Department of Head and Neck Surgery, University College London Hospital; †Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS), University College London; ‡OMFS specialist training, London Deanery; §Department of Medical Physics, University College London Hospital; ||Department of Craniofacial Surgery, Great Ormond Street Hospital; ¶Head and Neck Academic Centre, University College London

Disclosure: The authors declare that they have nothing to disclose.

S.C. is a Levi Fellow. This is a charitable salary funding fellowship based at University College London Hospital. This research was funded in whole, or in part, by the Wellcome Trust [203145Z/16/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright license to any author accepted article version arising from this submission.

S.C.: participated in research design, writing of the article, performance of the research, data analysis; A.T.: participated in the writing of the article; M.L. and Y.C.: participated in the performance of the research; E.E.: participated in research design and writing of the article; M.C. and C.S.: participated in research design, writing of the article, and data analysis.

Reprints: Soudeh Chegini, BMBCh, BDS, MA, Med, FRCS(OMFS), Clinical Research Fellow, University College London Hospital, Head and Neck Surgery, 1st Floor, West, 250 Euston Rd, London NW1 2PG, United Kingdom. Email: Soudeh.chegini@gmail.com and PhD Candidate, University College London, Charles Bell House, 43-45 Foley St, London W1W 7TY, United Kingdom

Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Annals of Surgery Open (2023) 2:e275

Received: 24 March 2022; Accepted 27 February 2023

Published online 16 May, 2023

DOI: 10.1097/AS9.0000000000000275

the tool in clinical practice. The learners' models and the time taken to produce the models are measured as a surrogate for learning gained using the training tool. The perception and value of the tool are evaluated using a postintervention questionnaire.

MATERIAL AND METHODS

Ethical approval was granted by the University College London ethics committee (reference number 19597/001). The study was open to professionally registered dental and medical clinicians aged 25–65 years, recruited via a social media platform.

TOOL DEVELOPMENT

For this project, a specific training tool was designed to be applied to the open-access 3Dslicer software.⁸ The tool constitutes structured didactic material, prescriptive exercises, and a postintervention questionnaire. Learners were also supplied with access to download the software and 3 computerized tomogram (CT) Digital Imaging and Communications in Medicine (DICOM) files (constitutes 6 fibulas). DICOM is the standard format medical scan that can be downloaded and transferred. Each scan contained a left and right fibula within the body of the scan. These were all arterial phase CT angiograms of the lower limbs of 0.75 mm³ voxel dimension. All scans were acquired from embodi3D, an online repository of open-access scans for educational and research purposes.

The didactic material was delivered in the form of a written manual, a prerecorded video, and a live video tutorial. All learners complete all 3 modalities. A blended approach was selected to meet the diverse learning approaches of adult learners and the varied time constraints of busy professionals. This approach was congruent with COVID restrictions.

The didactic material acted as a manual for the segmentation of the first CT scan to produce a 3D fibula model. The production of a fibula bone model was chosen as the example learning case. A bone model was selected due to its distinctive radio opacity. A single-segment model (bone only) is the simplest process for novices to follow while learning how to use the software.

The tool consists of a series of steps that instruct the learner to produce a fibula model from CT DICOM.⁹ Each step is aligned with a 'screen capture' which acts as a complementary stage setting picture.¹⁰ The procedural element provides rigid rules and limits decision-making. Discretionary judgment can be difficult and distracting for learners starting from the novice stage as acknowledged in the Dreyfus-Dreyfus model of skill acquisition.⁷

Subsequently, the learners were given a series of exercises, during which the learners were asked to repeat the steps to produce 5 further fibula models. This was completed in their own time with access to the didactic material and any other notes and sources of their own choosing. Online access to the instructor was available on request. It has been shown that prescriptive exercises have a higher completion rate than the 'on-you-own' exploration of software.¹¹

The didactic material included declarative commentary alongside the procedural steps. This was relevant on re-reading of the material to complete the exercises. The declarative element provides the context within the broader learning of the software abilities and purpose of the task. This holistic view allows the learner to prioritize important aspects and perceive deviation from normal patterns. In this way, the didactic material was designed to accommodate for both the novice stage of skill acquisition and the transition to proficiency.⁷

DATA ANALYSIS

Two medical physics technicians who regularly produce 3D models were also asked to produce fibula models from the same scans. One technician had 16 years of experience and

used Robins' 3D software. The second technician has 8 years of experience across clinical and academic image segmentation and used Mimics 3.0 from Materialise in this project. The technicians work at 2 separate large teaching National Health Service hospitals and produced the models for this study independent of each other. They did not have access to the training tool evaluated in this study. These models were used as the quality control, assuming that they would be of the standard used in regular clinical practice.

Learners submitted the models, as digital files, as they were produced along with the time taken to produce. For analysis, the Standard Tessellation Language models were imported into Meshlab.¹² The models were then compared using Hausdorff distance analysis. This is an automated analysis tool within the Meshlab software. Hausdorff distance analysis measures how far 2 subsets of a metric space are from each other. In essence, it utilizes the distance between multiple points on the surface of two 3D models. It is commonly used to compare 3D models or objects.

However, 18,925 measurement points were used for all Hausdorff distance measurements. Each model pair was measured twice, alternating the base model for the Hausdorff distance measurement. The reported Hausdorff distance measurements in this study are the average between these 2 readings. All statistics were calculated using SPSS version 27.

POST-INTERVENTION QUESTIONNAIRE

An end-of-course questionnaire was designed to assess the subjective opinion of learners on the value of this tool and its application to clinical practice. This was tested by a series of connected open questions (box 1) which starts with a closed binomial question, then an opinion on own practice, and then an opinion on the practice of the wider surgical community. This is based on the apprehension, exploration, cooperation, participation interview model.¹³ Thematic analysis was applied to analyze the questions 2–5.¹⁴

Box 1: Post-intervention Questionnaire items

1. Would you use this software again?
2. What area of practice can you apply this software?
3. And how will this improve your practice? If not, why?
4. What are the benefits of this course for other surgeons?
5. What are the limitations of the applications of this course for other surgeons?

RESULTS

In a 1-month period, 60 persons responded to the study recruitment. Of these 32 consented to participation. Due to study time and staff limitations, 10 learners were invited to undertake the course. All 10 completed the study producing 6 fibula 3D models for inclusion in this analysis.

All participants were between the ages of 25–40 years, 5 males and 5 females. All but 1 participant practiced as experienced oral and maxillofacial surgery junior doctors with experience in fibula surgery and 3D printing. None of the participants had any experience with medical image segmentation.

VARIABILITY BETWEEN TECHNICIANS

Figure 1 shows the variation in paired models produced by the 2 experienced technicians. The Hausdorff distance stands at less than 1 mm for all models. The second model has the highest Hausdorff distance value suggesting it was the most subjective model to segment. Table 1 and Fig. 2 show the time taken to produce models. It can be seen that the mean time taken to produce models decreases over time. A Spearman's correlation was run to determine the relationship between the time taken to produce each model and the order the model was produced. There was a negative monotonic

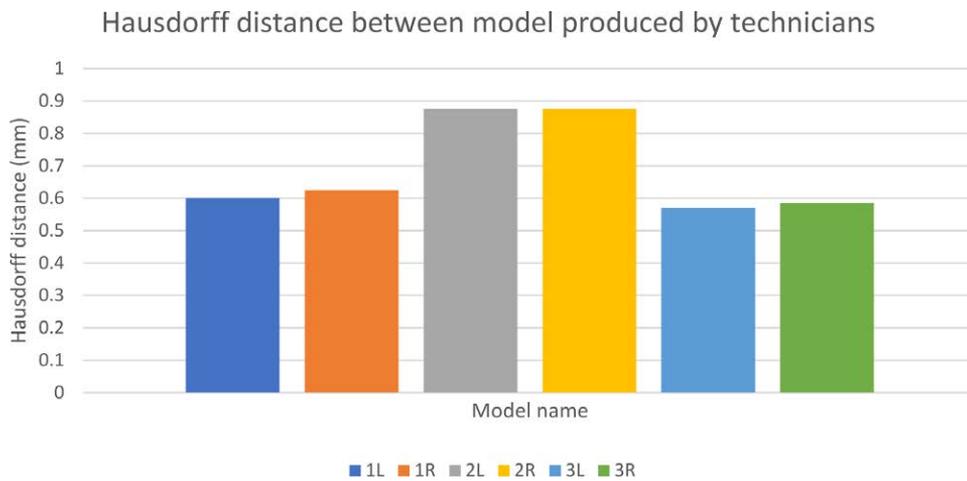


FIGURE 1. Hausdorff distance between models produced by the technician. Time taken to produce each sequential model

TABLE 1.
Time Taken to Produce Model in Order of Production

Order	Mean (min)	N	Std. Deviation (min)	Std. Error of Mean (min)	Minimum (min)	Maximum (min)	Range (min)
1	45.63	10.00	29.40	9.28	21.37	118.00	96.63
2	28.47	10.00	29.73	9.40	6.00	110.00	104.00
3	23.67	10.00	25.52	8.07	4.00	91.00	87.00
4	20.78	10.00	20.57	6.50	4.00	71.00	67.00
5	17.92	10.00	12.18	3.85	5.00	50.00	45.00
6	16.07	10.00	11.45	3.62	5.00	46.00	41.00
Total	25.42	60.00	23.92	3.08	4.00	118.00	114.00

Time taken to complete sequential models for each participant

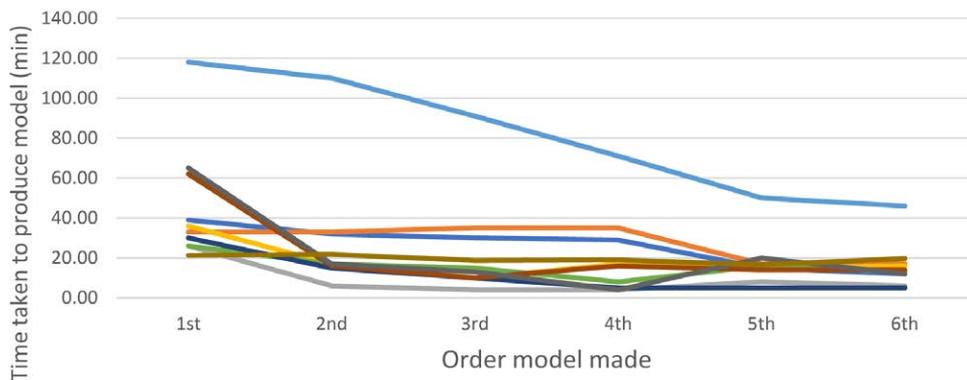


FIGURE 2. Time taken to produce sequential models. Each color represents a learner. Proportion is valued as 1 being same time taken to produce 1st model. Values less than 1 represent less time taken to produce a model compared to time taken to produce 1st model. Hausdorff distance between model produced by learners and technician.

correlation ($r_s = -0.375$; $n = 60$; $P = 0.003$). One learner failed to achieve a sub 30 minute on the 6th iteration. They were offered further practice and were able to achieve this target with their 8th-produced model.

Figure 3 shows the mean and range of Hausdorff distance variation by learners. Model 2 shows the greatest model variation. Model 3 which was produced last by learners shows the least model variation.

LOCALISATION OF VARIATION

Figure 4 is taken from meshlab. The 2L models made by learners and technicians are all superimposed. The heatmap shows

the regions of greatest variation between the models (blue/green). The is shown to be at the fibula ends and blood vessel attachment.

QUESTIONNAIRE

In response to question 1, all learners reported finding the training tool useful and would use the acquired learning in future clinical practice. Several key themes emerged during the thematic analysis of the qualitative data generated through the questionnaire: Current applications, future impact/benefits, and modeling skills. These are shown in Table 2.

Downloaded from http://journals.lww.com/annalsopen by BihDMf5ePpkav1ZEoum1iQIN4a+kLLNEZqpsIH04XIM0hCwWCX 1AWnYQp/1QHID3i3D0OQRy17VStf4C3VC4/OAVpDDa8KKGKGV0Ymy+78= on 06/28/2023

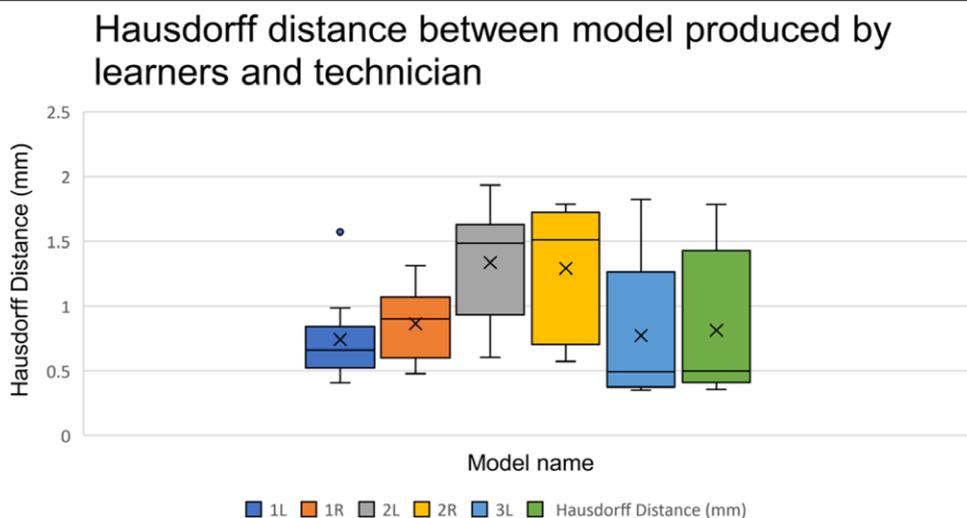


FIGURE 3. Boxplot showing Hausdorff distance between model produced by learners and technician

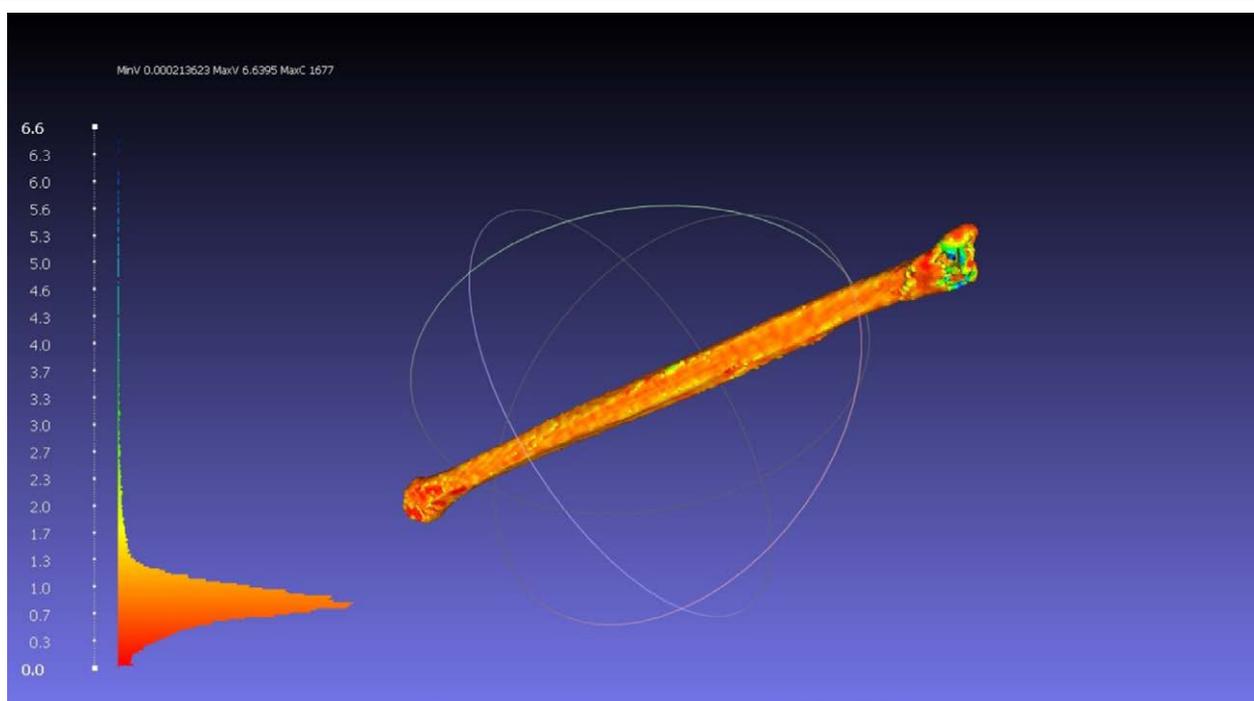


FIGURE 4. Image taken from meshlab computer software showing the superposition of 2L fibula models. The heatmap shows the areas which are most (green, yellow) and least similar (orange, red) between the models. The histogram on the left shows that red areas have the least difference - nearing 0mm difference. The fibula end and halfway along the shaft where blood vessels are attached to the bone show the most difference in segmentation.

CURRENT APPLICATIONS

Learners recognized the benefit of this training both within and beyond their own specialties, listing a variety of procedures. Making their own models was frequently reported as being beneficial to treatment planning.

“It allows us to provide a visual assessment before treatment.” (learner 10)

“This software can be used for planning three-dimensionally by other specialties as well.” (learner 4)

FUTURE IMPACT/BENEFITS

The enhancement to treatment planning was perceived to be associated with improved clinical outcomes, efficient use of resources, and as a communication tool with other professionals

and patients. Learners also recognized the cost saving of making their own models compared to outside technical services.

“... improving patient outcomes and optimizing time” (learner 1).

“A software with relatively easy learning curve as this might be time-saving, useful in eliminating the need for third-party help, and any surprise elements one might come across intraoperatively.” (learner 3).

MODELLING SKILLS

Making models can be time-consuming and not suitable for all clinicians. Despite this reality, learners reported many benefits of learning how models are made. This included understanding the limits of this technology, learning terminology, and appreciating

Downloaded from http://journals.lww.com/laosopen by BMDMf5ePpkav1ZEoum11QIN4a+kLLNEZqpsIH04XM0hCwCX 1AWnYQp/1QIH3D00QRy17vSF4Cf3V/C4/OAVpDDa8KKGKv0Ymy+78= on 06/28/2023

TABLE 2.
Frequency of excerpts representing identified codes in the questionnaire responses

Code	Excerpt Frequency	Theme
Treatment planning	6	Current applications
Saving time	5	Future impact/ benefits
Reducing healthcare costs	1	Future impact/ benefits
Easy-to-use software	1	Modeling skills
Eliminating the need for external companies	2	Future impact/ benefits
Visualizing anatomical structures	4	Future impact/ benefits
Improving patient outcomes	3	Future impact/ benefits
Software understanding and terminology	1	Modeling skills
Multi-disciplinary team communication	1	Future impact/ benefits
Teaching	1	Future impact/ benefits
Errors in segmentation	2	Modeling skills

the variations introduced in the process and in what clinical cases the variation would be unacceptable.

“I have gained some understanding of software option and general terminology” (learner 6).

“Using this software may not be reproducible because the segmentation of the tumor can be subjective.” (learner 5).

DISCUSSION

The ability to expertly perform a task as a learnt series of steps has been likened to being a skilled technician. However, a true expert should be able to tackle novel problems beyond the learnt procedure steps. This is known as adaptive expertise and is a better description of expertise among clinicians. Adaptive experts stretch the limits of their experience with flexibility and creativity. In this way, they can conquer novel and complex situations.⁴ This demands experience of how to perform every step of a task even if those elements are ultimately delegated. During medical training, clinicians are taught how to perform many tasks which are delegated to the multi-disciplinary team. This includes tasks from performing an electrocardiography to setting up surgical equipment in the theatre.

The authors are not aware of any publication which evaluates a training tool for clinicians on how to make 3D models from medical imaging. The Radiological Society of North America has held courses on 3D model creation. However, these courses have not been evaluated. Second, they are based on the Mimics software which is not open-access and thus not available to all clinicians.¹⁵ Clinicians who wish to employ this technology are dependent on technicians and may not apprehend the limitations of this technology. In turn, the technicians do not appreciate the constraints of medical therapy for which they are designing the 3D models.

The authors appreciate that 3Dslicer does not hold United States Food and Drug Administration or Conformité Européenne approval.¹⁶ However, it is routinely used in research and teaching to make models. It offers a comparable ability to its approved counterpart and is an acceptable platform to learn the principles of image segmentation.⁸

The subjective nature of image segmentation is illustrated in the variation in Hausdorff distance between technicians in this study. Although small, this variation depends on the quality of medical imaging and the patient's anatomy. The regions of highest variation are where the structure of interest is close to other structures with the same imaging properties. The heat map of

variation (Fig. 4), demonstrates this to be the fibula bone ends. Appreciating the source of these variations and the impact on case selection was highlighted in the responses in the postintervention feedback.

The tool evaluated in this study was developed based on the published understanding of software educational material for novices. Grounding the tool in established knowledge allows the impact of the tool on the target audience to be evaluated rather than testing the elements within the tool itself. The format of the tool itself is not innovative. However, the development of such a tool for clinicians and measuring the impact on their learning and practice is novel.

The learning of the skill was measured by the time taken to produce the models and comparing the models to that produced by technicians. The data shows that learners time taken to produce models statistically reduced as they progressed through the course. Similarly, the models improved. Scan 2 was the first scan segmented by the learners during the ‘prescriptive exercises’. The exercises were performed without supervision. As expected, the analysis of model 2 showed the greatest variation.

The mean difference between the final models made by learners and the technicians was less than that between the technicians themselves. This supports the progression of the learners from novice to proficient as in the Dreyfus-Dreyfus model of skill acquisition.⁷ This model is commonly applied to medical education and follows the transition of learners from novice, competent, proficient, and expert to mastery. The initial levels require recollection which is represented by learners following the procedural steps within the tool manual. This is followed by recognition which is supported by the declarative text within the tool. The latter stages require decision-making and awareness. The achievement of these latter stages is supported by the improved final fibula models made by learners.

In the case of the fibula bone, there is variable radiolucency at the bone ends. Depending on the Hounsfield unit applied by the user, sections of the end can be missed and thus excluded from the final model. Another important area is the margins of the bone with other structures. This can be other bones such as the Tibia or blood vessels (in CT angiogram scans). If there is not a clear space between these structures, users have to apply their best judgment and select the plane to divide these structures. This illustrates the subjective nature of image segmentation and the introduction of variation in 3D model making. These regions give rise to the variation in the produced models between learners and technicians and the technicians among themselves.

Segmenting these regions requires decision-making and is subjective. The final models show that the average variation between models made by learners compared to technicians is less than the technicians among themselves. This is despite the significantly more experience of the technicians. Using the Dreyfus-Dreyfus model, it can be postulated that the learners have increased awareness and make improved decisions, achieving more accurate models. This can be attributed to the application of their knowledge of anatomy from medical practice.

The postintervention questionnaire was used to measure the learner perceived impact of the tool on practice. Analysis of the questionnaire showed all learners found the tool useful. Thematic analysis of responses found this to be in a wide range of surgical practices, both in surgical planning and specific procedures. There was also a perceived benefit for nonsurgical practice including communication with the team and patients, teaching, and research.

This study presents a novel training tool, with high learner satisfaction, to teach clinicians to produce 3D models quickly and comparable models to technicians. Learners perceived the application of this skill to improve surgical planning, communication and teaching, their understanding of digital modeling, and improve clinical outcomes.

ACKNOWLEDGMENTS

The authors would like to thank the funding bodies who made this work possible. The first author is a Levi Fellow, this is a charitable fellowship based at University College London Hospital which provided salary funding. This research was funded in whole, or in part, by the Wellcome Trust [203145Z/16/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

REFERENCES

1. Zein NN, Hanouneh IA, Bishop PD, et al. Three-dimensional print of a liver for preoperative planning in living donor liver transplantation. *Liver Transpl.* 2013;19:1304–1310.
2. Diment LE, Thompson MS, Bergmann JHM. Clinical efficacy and effectiveness of 3D printing: a systematic review. *BMJ Open.* 2017;7:e016891.
3. Waran V, Narayanan V, Karupiah R, et al. Utility of multimaterial 3D printers in creating models with pathological entities to enhance the training experience of neurosurgeons. *J Neurosurg.* 2014;120:489–492.
4. Mylopoulos M, Regehr G. Cognitive metaphors of expertise and knowledge: prospects and limitations for medical education. *Med Educ.* 2007;41:1159–1165.
5. McCulloch P. The IDEAL framework for ensuring safety and effectiveness of medical devices. *BMJ.* 2020;370:m3183.
6. Van Der Meij H KJ, Stehouder M. Three decades of research and professional practice on printed software tutorials for novices. *Tech Commun (Washington).* 2009;56:265–292.
7. Dreyfus SE, Dreyfus H. a five-stage model of the mental activities involved in directed skill acquisition; 1980.
8. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging.* 2012;30:1323–1341.
9. Farkas DK. The logical and rhetorical construction of procedural discourse. *Tech Commun (Washington).* 1999;46:42–54.
10. Price J. How to write a computer manual. A handbook of software documentation. Penjamin/Cummings; 1984.
11. Wiedenbeck SZ, Nawyn J. An activity-based analysis of hands-on practice methods. *J Comput Assist Learn.* 2000;16:358–365.
12. Cignoni MC P, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G, MeshLab: an Open-Source Mesh Processing Tool. Sixth Eurographics Italian Chapter Conference. 2008:129–36.
13. Spradley JP. The ethnographic interview. Holt, Rinehart and Winston; 1979.
14. Braun V, Clarke V. What can “thematic analysis” offer health and wellbeing researchers? *Int J Qual Stud Health Well-being.* 2014;9:26152.
15. Wake N, Alexander AE, Christensen AM, et al. Creating patient-specific anatomical models for 3D printing and AR/VR: a supplement for the 2018 Radiological Society of North America (RSNA) hands-on course. *3D Print Med.* 2019;5:17.
16. Di Prima M, Coburn J, Hwang D, et al. Additively manufactured medical products - the FDA perspective. *3D Print Med.* 2016;2:1.