The accuracy of clinical 3D printing in reconstructive surgery: literature review and in vivo validation study

Michael P. Chae¹,²,³, Ru Dee Chung¹,²,³, Julian A. Smith³, David J. Hunter-Smith¹,²,³, Warren Matthew Rozen¹,²,³[^]

¹Department of Plastic, Reconstructive and Hand Surgery, Peninsula Health, Frankston, Victoria, Australia; ²Peninsula Clinical School, Central Clinical School at Monash University, The Alfred Centre, Melbourne, Victoria, Australia; ³Department of Surgery, School of Clinical Sciences at Monash University, Monash Medical Centre, Clayton, Victoria, Australia

**Contributions:** (I) Conception and design: All authors; (II) Administrative support: DJ Hunter-Smith, JA Smith; (III) Provision of study materials or patients: DJ Hunter-Smith, WM Rozen; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Prof Warren Matthew Rozen, MBBS, BMedSc, MD, PhD, FRACS. Department of Plastic, Reconstructive and Hand Surgery, Peninsula Health, PO Box 52, Frankston, Victoria 3199, Australia. Email: warrenrozen@hotmail.com.

**Abstract:** A growing number of studies demonstrate the benefits of 3D printing in improving surgical efficiency and subsequently clinical outcomes. However, the number of studies evaluating the accuracy of 3D printing techniques remains scarce. All publications appraising the accuracy of 3D printing between 1950 and 2018 were reviewed using well-established databases, including PubMed, Medline, Web of Science and Embase. An in vivo validation study of our 3D printing technique was undertaken using unprocessed chicken radius bones (*Gallus gallus domesticus*). Calculating its maximum length, we compared the measurements from computed tomography (CT) scans (CT group), image segmentation (SEG group) and 3D-printed (3DP) models (3DP group). Twenty-eight comparison studies in 19 papers have been identified. Published mean error of CT-based 3D printing techniques were 0.46 mm (1.06%) in stereolithography, 1.05 mm (1.78%) in binder jet technology, 0.72 mm (0.82%) in PolyJet technique, 0.20 mm (0.95%) in fused filament fabrication (FFF) and 0.72 mm (1.25%) in selective laser sintering (SLS). In the current in vivo validation study, mean errors were 0.34 mm (0.86%) in CT group, 1.02 mm (2.51%) in SEG group and 1.16 mm (2.84%) in 3DP group. Our Peninsula 3D printing technique using a FFF 3D printer thus produced accuracy similar to the published studies (1.16 mm, 2.84%). There was a statistically significant difference (P<10^-4) between the CT group and the latter SEG and 3DP groups indicating that most of the error is introduced during image segmentation stage.

**Keywords:** Accuracy; validation; 3D printing; in vivo; plastic surgery

Submitted Apr 20, 2021. Accepted for publication Jun 23, 2021.

doi: 10.21037/gs-21-264

**View this article at:** https://dx.doi.org/10.21037/gs-21-264

**Introduction**

A growing number of studies demonstrate the utility of 3D printing for preoperative planning, creating intraoperative guidance devices, improving patient education and medical training, and designing patient-specific implants (1-3). Moreover, 3D printing may improve surgical efficiency and accuracy (4), leading to reduced operating time and subsequent health economic benefits (5). However, 3D printing has not yet been incorporated as part of routine clinical practice due to several significant shortcomings (6).

The main limitations of 3D printing are high cost, lengthy production time and perceived inaccuracy of the models (6). Recently, the cost and production time of 3D printing has decreased due to technological advances, and 3D printing is being widely adopted in various fields of medicine.
Accurate medical 3D printing have reduced significantly by the introduction of affordable 3D printers and free software suites (2,7,8). Encouraged by this, Chae et al. have published the Peninsula 3D printing technique using a desktop 3D printer and an open-source 3D software (2). However, there is a relative paucity of studies investigating the accuracy of 3D printing in the literature (9).

Error can be introduced at any stage of 3D printing, but mainly during image acquisition and segmentation (Figure 1) (6,10-12). Most studies have examined the accuracy of 3D printing techniques by comparing linear measurements of the original object to the final 3D-printed model (13-31).

The current study reviews of all published studies where accuracy of image-guided 3D printing of cadaveric specimens is evaluated. Subsequently, an in vivo validation study of the published Peninsula 3D printing technique (2) using the chicken radius bone (Gallus gallus domesticus) is conducted.

Methods

Literature review

A review was undertaken reviewing the published English literature from 1950 to 2018 from well-established databases, such as PubMed, Medline, Web of Science and Embase using search terms, such as “3D printing”, “additive manufacturing”, “rapid prototyping”, “stereolithography”, “CT”, “MRI”, “accuracy”, “precision”, “validation”, “evaluation” and “comparison”. Secondary references found through bibliographic linkage were also retrieved.

Inclusion criteria

Only studies using cadaveric specimens where their direct measurements could be obtained as gold standard were included. Anatomical structures were chosen over generically-shaped phantoms since they pose greater challenge in dimensional assessment and the findings will be more relevant to clinical application. Only studies using CT as imaging source were included since an overwhelming majority of medical 3D printing has been conducted using it (9).

Exclusion criteria

Studies were excluded if authors indirectly deduce the accuracy of 3D printing by analysing the clinical placement of implants or screws using 3D-printed surgical guides since this introduces potentially confounding clinical variables (32-38). Studies using qualitative (39) or volumetric assessment (40), instead of linear measurements, to report 3D printing accuracy were excluded (41-43).

Validation study

The chicken radius bone was chosen for this study, instead of other bones such as the humerus, femur, tibia or fibula, since its length and shape closely resembles the human proximal phalanx (range of 37–44 vs. 32–45 mm, respectively) (Figure 2) (44). Furthermore, it is readily available from the butchers intact in large numbers, compared to other parts of the chicken that are normally discarded or damaged during preparation. This validation study was approved by the Human Research Ethics Committee at Peninsula Health.

Figure 1 Steps involved in creating a CT-based 3D-printed anatomical biomodel. Error can be introduced at each step. CT, computed tomography; DICOM, digital imaging and communications in medicine; STL, standard tessellation language.
**In vitro measurement**

The entire defeathered chicken wing was wrapped in cling wrap, placed inside a radiolucent polypropylene-based standard household container, and CT scanned. Then, the radius was dissected out and its articular cartilages removed. Maximum length of the bone was measured using a digital Vernier caliper (Kincrome, Scoresby, VIC, Australia). Out of 50 chicken wings scanned, 11 were removed from the study due to detection of their unrecognised midshaft fractures from the CT scan, which can compromise reliability of the caliper measurements.

**CT image acquisition**

CT scanning was performed using Siemens SOMATOM Definition AS scanner (Siemens AG, Erlangen, Germany) and hand/wrist scanning protocol. The scan parameters were as follows: tube voltage of 120 kV, tube current exposure time product of 35 mA, single collimation width of 0.6 mm and slice thickness of 0.6 mm. The images were then reconstructed using the bone window and exported in Digital Imaging and COmmunication in Medicine (DICOM) format (1,054 images). The length of each chicken radius was measured from the CT scan using OsiriX software (Pixmeo, Geneva, Switzerland) and recorded to two decimal places (Figure 2).

**Image segmentation**

The DICOM files were reconstructed into 3D images using free, open-source 3D Slicer software (Surgical Planning Laboratory, Boston, MA, USA) as previously reported and exported in standard tessellation language (STL) format (2,45,46). Briefly, using “ThresholdEffect” function, a range was set from the Hounsfield unit-derived values in order to automatically segment chicken radius from the surrounding soft tissue. Using “PaintEffect” tool, minor surface artefacts were manually removed. Minor surface artefacts were defined as irregularities that did not change overall structure or shape, and which not associated with continuous lines or borders. Each radius was labelled and exported separately. The length of each bone was measured from its 3D-reconstructed image, or STL file, using Autodesk MeshMixer software (Autodesk Inc., San Rafael, CA, USA).

**3D printing**

Each STL file was converted into a 3D printer-friendly G-code file using Cura software (Ultimaker, Geldermalsen, The Netherlands). The bones were 3D-printed in thermoplastic polylactic acid (PLA) filament using a desktop, fused filament fabrication (FFF) Ultimaker 3 Extended 3D printer (Ultimaker, Geldermalsen, The Netherlands).
Netherlands).

Statistical analysis

All measurements were repeated three times and the average was recorded in millimetres to two decimal places. Comparative analysis was performed using Stata statistical software package (StataCorp, College Station, TX, USA). The mean lengths were analysed using one-way ANOVA technique. Using direct caliper measurements as the gold standard, the difference was measured and analysed using unpaired two-sample t-test with unequal variances, with Welch approximation. P value of less than 0.05 was accepted as statistically significant.

Results

Literature review

Characteristics of the published 3D printing validation studies

There were 28 comparison studies within 19 papers which examined the accuracy of medical 3D printing between 1994 and 2016 (Table 1 and Figure 3). Majority of investigators use multidetector CT (MDCT) over cone-beam CT (CBCT) as imaging modality (24 vs. 4 studies). Cadaveric skull and mandible are most commonly used as specimens (16 and 10 studies, respectively). This is most likely due to the fact that 3D printing has been most extensively studied for application in cranio-maxillofacial surgery. Relatively consistent quality of CT images has been acquired as evidenced by a narrow range of CT layer thickness used (0.5–2.5 mm). Majority of papers have reported 3D printing in-house, rather than outsourcing to an external company (14 vs. 5 papers). Encouragingly, this appears to be a growing trend (median year of publication: 2011.5 vs. 2009 respectively). Linear measurements were made more commonly using a caliper than automated CMM (22 vs. 6 studies).

Published accuracy of 3D printing

Stereolithography and binder jet technique (BJT) were most commonly investigated (10 and 7 studies, respectively), followed by newer selective laser sintering (SLS) (4 studies), FFF (4 studies) and PolyJet technology (PJT) (3 studies) (Table 1).

Stereolithography

Stereolithography (SLA) is the earliest 3D printing technology described and is considered the gold standard in 3D biomodel production (47). However, it is slow, relatively expensive and requires significant post-production manual processing (2). Most studies used MDCT over CBCT (8 vs. 2 studies) at slice thickness of 0.5–1.5 mm, and cadaveric skull and mandible as specimens (8 and 2 studies, respectively) (13-21). Highest z-resolution of SLA 3D printers ranged between 0.00177 and 0.0762 mm. Using MDCT, the mean absolute difference between the original specimen and its 3D-printed model was 0.46 mm (range, 0.12–0.85 mm) and relative difference 1.06% (0.08–2.74%). Using CBCT, the mean difference was 0.35–0.39 mm and 0.74–0.82%.

BJT

BJT, also known as powder bed technique, describes a process where a printer head ejects binder material and coloured dye simultaneously onto a bed of powder and fuses them layer-by-layer into a plaster model (48). Major benefits of BJT are that it forgoes support structures and can produce models in multiple colours and materials. However, the final product is brittle and requires extensive post-production manual processing (2). Similar to SLA, most studies used MDCT over CBCT (6 vs. 1 studies) at slice thickness of 0.5–1.5 mm, and cadaveric skull, mandible and upper limb bones as specimens (3, 3, and 1 studies, respectively) (22-25). Highest z-resolution of BJT 3D printers ranged between 0.0762 mm and 0.12 mm. Using MDCT, the mean difference was 1.05 mm (0.32–2.00 mm) and 1.78% (0.38–3.14%). Using CBCT, the mean difference was 0.63 mm and 1.87%.

PJT

PJT, also known as multijet modelling, is similar to SLA but the liquid photopolymer is immediately cured by the ultraviolet (UV) light preventing time-consuming post-production processing (49). A major benefit of PJT is its high resolution; however, its surface finish is still inferior to SLA and the printers remain expensive (2). All studies of PJT used MDCT at slice thickness of 0.6 mm, and cadaveric mandible and skull as specimens (2 and 1 studies, respectively) (23-25). Highest z-resolution of PJT 3D printers was 0.016 mm. Using MDCT, the mean difference was 0.72 mm (0.20–1.23 mm) and 0.82% (0.13–2.14%).

FFF

FFF is the most affordable consumer-grade 3D printing technology where a melted filament of thermoplastic material is extruded in a layer-by-layer fashion (50,51). Major advantages of FFF is its cost and convenience; however, it almost inevitably requires simultaneous
Table 1 Summary of studies that assess the accuracy of medical 3D printing using linear measurements on cadaveric specimens

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Specimen</th>
<th>Outsource (Y/N)</th>
<th>CT Type</th>
<th>Slice thickness (mm)</th>
<th>Name</th>
<th>3D Printing Company</th>
<th>Z-resolution (mm)</th>
<th>Name</th>
<th>3D printing software</th>
<th>Mode of measurement</th>
<th>Mean difference (mm)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLA Primo et al.</td>
<td>2012</td>
<td>Skull</td>
<td>MDCT</td>
<td>N/A</td>
<td>ZPrinter 310</td>
<td>3D Systems</td>
<td>0.0762</td>
<td>Magics</td>
<td>Materialise NV</td>
<td>Caliper</td>
<td>N/A</td>
<td>0.62</td>
<td>(13)</td>
</tr>
<tr>
<td>Taft et al.</td>
<td>2011</td>
<td>Skull</td>
<td>MDCT</td>
<td>0.625</td>
<td>Viper 2si</td>
<td>3D Systems</td>
<td>0.0625</td>
<td>Mimics</td>
<td>Materialise NV</td>
<td>CMM</td>
<td>N/A</td>
<td>0.61</td>
<td>(14)</td>
</tr>
<tr>
<td>Nizam et al.</td>
<td>2006</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.25</td>
<td>Viper 2si</td>
<td>3D Systems</td>
<td>0.0625</td>
<td>Mimics</td>
<td>Materialise NV</td>
<td>Caliper</td>
<td>0.23</td>
<td>0.08</td>
<td>(15)</td>
</tr>
<tr>
<td>Choi et al.</td>
<td>2002</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.0</td>
<td>SLA 5000</td>
<td>3D Systems</td>
<td>0.0177</td>
<td>V-Works</td>
<td>CyberMed</td>
<td>Caliper</td>
<td>0.62</td>
<td>0.56</td>
<td>(16)</td>
</tr>
<tr>
<td>Asaumi et al.</td>
<td>2001</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.0</td>
<td>N/A</td>
<td>D-MEC</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>N/A</td>
<td>0.63</td>
<td>(17)</td>
</tr>
<tr>
<td>Bouyssié et al.</td>
<td>1997</td>
<td>Mandible</td>
<td>MDCT</td>
<td>1.0</td>
<td>SLA 250</td>
<td>3D Systems</td>
<td>0.0625</td>
<td>Mimics</td>
<td>Materialise NV</td>
<td>Caliper</td>
<td>0.12</td>
<td>2.10</td>
<td>(18)</td>
</tr>
<tr>
<td>Ozo et al.</td>
<td>1994</td>
<td>Skull</td>
<td>MDCT</td>
<td>0.5</td>
<td>EDS-1000 HD</td>
<td>D-MEC</td>
<td>0.02</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>0.33</td>
<td>0.68</td>
<td>(19)</td>
</tr>
<tr>
<td>Barker et al.</td>
<td>1994</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.5</td>
<td>SLA 250</td>
<td>3D Systems</td>
<td>0.0625</td>
<td>Analyze</td>
<td>AnalyzeDirect</td>
<td>Caliper</td>
<td>0.85</td>
<td>2.74</td>
<td>(20)</td>
</tr>
<tr>
<td>Primo et al.</td>
<td>2012</td>
<td>Skull</td>
<td>CBCT</td>
<td>N/A</td>
<td>ZPrinter 310</td>
<td>3D Systems</td>
<td>0.0762</td>
<td>Magics</td>
<td>Materialise NV</td>
<td>Caliper</td>
<td>N/A</td>
<td>0.74-0.82</td>
<td>(13)</td>
</tr>
<tr>
<td>Santana et al.</td>
<td>2012</td>
<td>Mandible</td>
<td>MDCT</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>0.35-0.39</td>
<td>N/A</td>
<td>(21)</td>
</tr>
<tr>
<td>BJT McMenamin et al.</td>
<td>2014</td>
<td>Upper Limb</td>
<td>MDCT</td>
<td>1.0</td>
<td>ZPrinter 650</td>
<td>3D Systems</td>
<td>0.089</td>
<td>Rhinoceros 3D</td>
<td>Robert McNeal &amp; Associates</td>
<td>Caliper</td>
<td>0.32</td>
<td>1.29</td>
<td>(22)</td>
</tr>
<tr>
<td>Salmi et al.</td>
<td>2013</td>
<td>Skull</td>
<td>MDCT</td>
<td>N/A</td>
<td>ZPrinter 450</td>
<td>3D Systems</td>
<td>0.089</td>
<td>Zeiss Calypso</td>
<td>Carl Zeiss</td>
<td>CMM</td>
<td>0.44-0.80</td>
<td>0.38-0.69</td>
<td>(23)</td>
</tr>
<tr>
<td>Murugesan et al.</td>
<td>2012</td>
<td>Mandible</td>
<td>MDCT</td>
<td>0.6</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>N/A</td>
<td>1.67</td>
<td>(24)</td>
</tr>
<tr>
<td>Ibrahim et al.</td>
<td>2009</td>
<td>Mandible</td>
<td>MDCT</td>
<td>N/A</td>
<td>ZPrinter 310</td>
<td>3D Systems</td>
<td>0.0762</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>1.44</td>
<td>3.14</td>
<td>(25)</td>
</tr>
<tr>
<td>Silva et al.</td>
<td>2008</td>
<td>Skull</td>
<td>YMDCT</td>
<td>0.5</td>
<td>ZPrinter 310</td>
<td>3D Systems</td>
<td>0.0762</td>
<td>InVesalius</td>
<td>CTI</td>
<td>Caliper</td>
<td>1.07</td>
<td>2.10</td>
<td>(26)</td>
</tr>
<tr>
<td>Change et al.</td>
<td>2003</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.5</td>
<td>ZPrinter 420</td>
<td>3D Systems</td>
<td>0.0762</td>
<td>Maximil</td>
<td>Medicim</td>
<td>Caliper</td>
<td>&lt;2.00</td>
<td>2.10-4.70</td>
<td>(27)</td>
</tr>
<tr>
<td>Oliszewski et al.</td>
<td>2014</td>
<td>Mandible</td>
<td>CBCT</td>
<td>0.5</td>
<td>Matrix 300+</td>
<td>Moc Technologies</td>
<td>0.1</td>
<td>Zeiss Calypso</td>
<td>Carl Zeiss</td>
<td>CMM</td>
<td>0.36</td>
<td>1.87</td>
<td>(28)</td>
</tr>
<tr>
<td>SLS Petrudis et al.</td>
<td>2015</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.0</td>
<td>EOSINT P 395</td>
<td>EOS GmbH</td>
<td>0.06</td>
<td>OtsiX</td>
<td>Pimeo</td>
<td>Caliper</td>
<td>0.16</td>
<td>0.30</td>
<td>(29)</td>
</tr>
<tr>
<td>Salmi et al.</td>
<td>2013</td>
<td>Skull</td>
<td>MDCT</td>
<td>N/A</td>
<td>EOSINT P 800</td>
<td>EOS GmbH</td>
<td>0.12</td>
<td>Zeiss Calypso</td>
<td>Carl Zeiss</td>
<td>CMM</td>
<td>0.93</td>
<td>0.79</td>
<td>(23)</td>
</tr>
<tr>
<td>Ibrahim et al.</td>
<td>2009</td>
<td>Mandible</td>
<td>MDCT</td>
<td>N/A</td>
<td>DTM Sinterstation 2000</td>
<td>3D Systems</td>
<td>0.254</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>0.90</td>
<td>1.79</td>
<td>(25)</td>
</tr>
<tr>
<td>Silva et al.</td>
<td>2008</td>
<td>Skull</td>
<td>MDCT</td>
<td>0.5</td>
<td>DTM Sinterstation 2000</td>
<td>3D Systems</td>
<td>0.254</td>
<td>InVesalius</td>
<td>CTI</td>
<td>Caliper</td>
<td>0.89</td>
<td>2.10</td>
<td>(26)</td>
</tr>
<tr>
<td>FFF Oglyen et al.</td>
<td>2015</td>
<td>Vertebra</td>
<td>MDCT</td>
<td>0.625</td>
<td>RepliCAD 2</td>
<td>MakerBot</td>
<td>0.1</td>
<td>Analyze</td>
<td>AnalyzeDirect</td>
<td>Caliper</td>
<td>0.18</td>
<td>0.69</td>
<td>(30)</td>
</tr>
<tr>
<td>Petrudis et al.</td>
<td>2015</td>
<td>Skull</td>
<td>MDCT</td>
<td>1.0</td>
<td>Cubix</td>
<td>3D Systems</td>
<td>0.1</td>
<td>OtsiX</td>
<td>Pimeo</td>
<td>Caliper</td>
<td>0.21</td>
<td>0.44</td>
<td>(29)</td>
</tr>
<tr>
<td>Murugesan et al.</td>
<td>2012</td>
<td>Mandible</td>
<td>MDCT</td>
<td>0.6</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>1.57</td>
<td>1.73</td>
<td>(24)</td>
</tr>
<tr>
<td>Maschio et al.</td>
<td>2016</td>
<td>Mandible</td>
<td>CBCT</td>
<td>0.5</td>
<td>Up Plus 2</td>
<td>Beijing TierTime Technology</td>
<td>0.15</td>
<td>Maximil</td>
<td>Medicim</td>
<td>CMM</td>
<td>0.37</td>
<td>3.76</td>
<td>(31)</td>
</tr>
<tr>
<td>PJT Salmi et al.</td>
<td>2013</td>
<td>Skull</td>
<td>MDCT</td>
<td>N/A</td>
<td>Objet Eden 350V</td>
<td>Stratasys</td>
<td>0.016</td>
<td>Zeiss Calypso</td>
<td>Carl Zeiss</td>
<td>CMM</td>
<td>0.20</td>
<td>0.18</td>
<td>(23)</td>
</tr>
<tr>
<td>Murugesan et al.</td>
<td>2012</td>
<td>Skull</td>
<td>MDCT</td>
<td>0.6</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>N/A</td>
<td>0.13</td>
<td>(24)</td>
</tr>
<tr>
<td>Ibrahim et al.</td>
<td>2009</td>
<td>Mandible</td>
<td>MDCT</td>
<td>N/A</td>
<td>Objet Eden 330</td>
<td>Stratasys</td>
<td>0.016</td>
<td>N/A</td>
<td>N/A</td>
<td>Caliper</td>
<td>1.23</td>
<td>2.14</td>
<td>(25)</td>
</tr>
</tbody>
</table>

CT, computed tomography; MDCT, multidetector CT; CBCT, cone beam CT; mm, millimetre; Y, yes; N, no; SLA, stereolithography; BJT, binder jet technique; SLS, selective laser sintering; FFF, fused filament fabrication; PJT, PolylJet technology; N/A, not available; CMM, coordinate measuring machine.
production of support structures and hence, yields the lowest quality of surface finish (2). Most studies used MDCT over CBCT (3 vs. 1 studies) at slice thickness of 0.5–1.0 mm, and cadaveric mandible, skull, and vertebra as specimens (2, 1, and 1 studies, respectively) (24,29-31).

Highest z-resolution of FFF 3D printers ranged between 0.1 mm and 0.15 mm. Using MDCT, the mean difference was 0.20 mm (0.18–0.21 mm) and 0.95% (0.44–1.73%). Using CBCT, the mean difference was 0.37 mm and 3.76%.

SLS
SLS is a technique where powdered forms of thermoplastic, metal, glass or ceramic material are sintered by high-power laser beams in a layer-by-layer manner (52). Similar to BJT, SLS foregoes support structures and is capable of producing delicate structures with smooth surface finishes. However, high-powered laser of SLS 3D printers requires expert handling for safety and hence, they are related to high cost (2). In contrast to SLA and BJT, all studies of SLS used MDCT at slice thickness of 0.5–1.0 mm, and cadaveric skull and mandible and upper limb bones as specimens (3 and 1 studies, respectively) (23,25,26,29). Highest z-resolution of SLS 3D printers ranged between 0.06 mm and 0.254 mm. Using MDCT, the mean difference was 0.72 mm (0.16–0.93 mm) and 1.25% (0.30–2.10%).

Validation study
Image segmentation took an average of 7.5 minutes per bone (4.2–13.4 minutes). 3D printing took an average of 28.8 minutes per model (25–37 minutes). Mean length of chicken radius bones from direct caliper measurements (i.e., the gold standard) was 39.54 mm (36.64–43.75 mm). Mean length from CT scans was 39.32 mm (36.10–43.30 mm), producing a difference of 0.22 mm from the gold standard (P=1.00). Mean length from image segmentation was 40.56 mm (37.18–44.96 mm), producing a difference of 1.02 mm from the gold standard (P=0.22). Mean length from 3D-printed models was 40.70 mm (37.17–45.00 mm), producing a difference of 1.16 mm from the gold standard (P=0.10).

In order to calculate mean absolute and relative difference, measurements from the DICOM images of CT scans (CT group), image segmentation (SEG group) and 3D-printed models (3DP group) are subtracted from the gold standard for each bone (Figure 3). Mean absolute difference from the CT group was 0.34 mm (standard deviation: 0.26–0.42 mm); SEG group: 1.02 mm (SD: 0.88–1.15 mm); and 3DP group: 1.16 mm (SD: 1.03–1.29 mm). Mean absolute difference from the CT group was significantly smaller than both

<table>
<thead>
<tr>
<th>Difference</th>
<th>CT</th>
<th>SEG</th>
<th>3DP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute (mm)</td>
<td>Mean 0.34</td>
<td>1.02</td>
<td>1.16</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.26–0.42</td>
<td>0.88–1.15</td>
<td>1.03–1.29</td>
</tr>
<tr>
<td>P&lt;10^−4</td>
<td></td>
<td>P=0.13</td>
<td></td>
</tr>
</tbody>
</table>

| Relative (%) | Mean 0.86  | 2.51         | 2.84         |
| 95% CI      | 0.66–1.06  | 2.17–2.85    | 2.54–3.15    |
| P<10^−4   |             | P=0.15       |              |

Figure 3 Summary of difference of the maximum length of chicken radius bone (gold standard) against CT DICOM images, SEG, and 3DP. CT, computed tomography; SEG, image segmentation; 3DP, 3D printing. *, P<0.05.
SEG (P<10^{-4}) and 3DP groups (P<10^{-4}). However, there was no statistical significance between the SEG and 3DP groups (P=0.13). Mean relative difference from CT group was 0.86% (SD: 0.66–1.06%); SEG group 2.51% (SD: 2.17–2.85%) and 3DP group 2.84% (2.54–3.15%). Similarly, mean relative difference of the CT group was significantly smaller than both SEG (P<10^{-4}) and 3DP groups (P<10^{-4}), but there was no statistical significance between latter two groups (P=0.15).

**Discussion**

Despite a plethora of studies demonstrating numerous utility of 3D printing in surgery (1-3), there is a relative paucity of studies assessing its accuracy in a systematic manner. Current review of 28 comparison studies in 19 papers demonstrates that all 3D printing techniques using conventional MDCT have acceptable accuracy for clinical application with mean difference of less than 1–2 mm (1–2%): SLA: 0.46 mm (1.06%), BJT: 1.05 mm (1.78%), PJT: 0.72 mm (0.82%), FFF: 0.20 mm (0.95%) and SLS 0.72 mm (1.25%). However, it is worth noting that studies have used varying approaches, making them difficult to compare directly.

Some investigators have used CBCT as imaging source, which has lower cost, lower radiation and higher spatial resolution for bony micro-architecture than MDCT, but it is prone to motion artefact and is limited by small field of view (average of 6 cm x 6 cm) making it ideal for dental imaging (53). A wide variety of software programs has been used for image segmentation that differ in basic algorithm for creating 3D-reconstructed images. Some software suites are difficult to obtain for widespread application due to their high cost, such as Mimics (Materialise NV, Leuven, Belgium), or are limited to their local institution, such as V-Works (CyberMed, Seoul, South Korea). Similarly, 3D printers of the same technique can differ in accuracy amongst different brands and different models within the same brand.

Errors can be introduced at all stages of medical 3D printing (Figure 1), but occurs most frequently and most significantly during image acquisition and segmentation (6,12,54). In comparison to MRI, CT is associated with less geometric distortion (55) and, therefore, greater accuracy (20), especially when 3D printing bones (56). The greatest discrepancy can occur from incorrect selection of CT slice thickness when exporting DICOM images (16). CT images are routinely acquired in slice thickness of 0.6–2.0 mm, ideal for 3D printing; however, these are only stored temporarily and discarded after being converted to thicker 5.0 mm images to reduce storage cost (57). As scanners continue to evolve, errors from other aspects of CT, such as image noise, beam hardening, motion artefact, metal artefact and gantry tilt distortion, have significantly improved (58).

When evaluating the accuracy of 3D printing, it is important to differentiate it from describing the z-resolution, which indicates the thinnest layer of 3D print that can be deposited per layer. However, this does not correlate with the overall accuracy of the final model, since increasing the number of layers that needs to be 3D-printed increases the potential for error and artefact (59).

Image segmentation involves partitioning DICOM images into multiple regions that correspond to an anatomical structure of interest and produces its digital 3D reconstruction (58). Due to its reliance on operator expertise for selecting appropriate threshold values and manual processing, image segmentation is highly susceptible to inter-operator variability and is potentially the most significant source of inaccuracy (11,12,60). In place of time-consuming traditional manual segmentation methods (61), semi-automated segmentation techniques have been developed, such as global thresholding (62), edge detection (63) and region growing (64). For 3D printing bones, global thresholding is the most commonly used technique (65,66) and is also employed in our validation study. Despite its benefits, it still requires additional manual editing (12,67). Recent improvements in artificial intelligence and machine learning capabilities pose interesting potential in this field (68). Furthermore, the size of each triangular mesh produced during 3D reconstruction, called triangulation algorithm, can impact accuracy of the final model (69).

Inaccuracy can still be introduced during post-segmentation image processing to further smooth or trim 3D images. However, this has improved significantly in recent times with software advancements and segmentation techniques (70). Slicing software and 3D printers are often packaged together and are generally considered precise by industry standards (71). The surface of 3D-printed models can be damaged during post-print processing to remove support structures, especially with FFF 3D printers. However, these superficial blemishes rarely significantly compromise the overall accuracy (58). Nonetheless, in our study, the models were 3D-printed in an upright position to minimise the amount of support structures that would
need to be removed. Altogether, the errors encountered at all the steps combine and influence each other, making the sum of all errors not additive, but multiplicative (9,25). As a result, when assessing 3D printing, the whole process from imaging to fabrication must be validated together.

The results from current validation study using chicken radius is comparable to previously published reports. Interestingly, amongst the studies using FFF 3D printers, our mean absolute and relative differences compare higher than the published values (1.16 vs. 0.20 mm and 2.84% vs. 0.95%, respectively). One of the reasons for this discrepancy may be from soft tissue artefact. In our study, the bones are scanned as chicken wings and then dissected out afterwards. In other studies, the cadaveric specimens are scanned after being dissected, which would yield superior contrast and spatial resolution at bone-to-air interface. Moreover, unlike other expensive, proprietary image segmentation programs, such as Mimics (Materialise NV, Leuven, Belgium), Vitrea (Vital Images, Minnetonka, MN, USA) and OsiriX MD (Pixmeo, Geneva, Switzerland), we have used 3D Slicer used, which has not yet been approved by the US Food and Drugs Administration (FDA) for routine clinical application, albeit only for research purposes (72).

Nevertheless, one of the major limitations of this validation study is that the findings cannot be directly translated to 3D printing anatomical structures that are different in size to chicken radius, such as human femur or humerus. The mean absolute difference found in our study may be altered with other bones in the body with different soft tissue artefacts and their unique anatomical characteristics affecting image segmentation. The mean relative difference may not be translatable since the degree of inaccuracy between different human bones may not linearly correlate with their size. This study is limited to commentary on the accuracy of 3D printing in small bones, and thus direct comparison to large bones is limited. Many of the sources of error, however, will exist in the case of small and large bones, and will exist whether done experimentally or in the clinical setting. In the experimental setting of this study, there is an inherent reduction of error by removing all soft tissue from the scanned bones, and a reduction in the motion artefact that would occur clinically, and thus it could be hypothesized that more of the error could be attributed to image segmentation.

**Conclusions**

There have been 28 comparison studies in 19 papers which demonstrate that all 3D printing techniques using conventional MDCT have acceptable accuracy for clinical application, with a mean difference from original anatomical specimens of less than 1–2 mm (1–2%): SLA: 0.46 mm (1.06%), BJT: 1.05 mm (1.78%), PJT: 0.72 mm (0.82%), FFF: 0.20 mm (0.95%), and SLS 0.72 mm (1.25%). The current validation study of the Peninsula 3D printing technique produced accuracy similar to the published studies (1.16 mm, 2.84%). There was a statistically significant difference (P<10⁻⁴) between the CT group and the latter SEG and 3DP groups indicating that most of the error is introduced during image segmentation stage.

**Acknowledgments**

We would like to acknowledge the help of Dr Vicky Tobin, our research coordinator, for her support with study design, data collection and analysis.

**Funding:** None.

**Footnote**

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at [https://dx.doi.org/10.21037/gs-21-264](https://dx.doi.org/10.21037/gs-21-264)). WMR serves as an unpaid editorial board member of *Gland Surgery* from Mar 2018 to Feb 2023. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: [https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/).

**References**


53. Posadzy M, Desimpel J, Vanhoenacker F. Cone beam CT of the musculoskeletal system: clinical applications.


