Evaluation of Water Displacement Method in Estimating Mandibular Ramus Autograft Volume

Ferit Bayram1, *, Senem Aşkı̇n Ekinci1, Gökhan Göçmen1

1 Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Marmara University, Istanbul, Türkiye

Corresponding Author:
Senem Aşkı̇n Ekinci, DDS
Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Marmara University, Istanbul, Türkiye
E-mail(s): askinesenem@gmail.com
ORCID id: https://orcid.org/0000-0002-4051-3172

Authors
Ferit Bayram, DDS, PhD, Assistant Professor
Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Marmara University, Istanbul, Türkiye
E-mail(s): ferit.bayram@marmara.edu.tr
ORCID id: https://orcid.org/0000-0002-4395-4547

Senem Aşkı̇n Ekinci, DDS
Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Marmara University, Istanbul, Türkiye
E-mail(s): askinesenem@gmail.com
ORCID id: https://orcid.org/0000-0002-4051-3172

Gökhan Göçmen, DDS, PhD, Associate Professor
Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Marmara University, Istanbul, Türkiye
E-mail(s): gokhan.gocmen@marmara.edu.tr
ORCID id: https://orcid.org/0000-0003-0317-4308

Article Type: Original Article

Acknowledgments: None

Informed Consent: All patients signed a written consent for academic use of their radiologic data.

Conflict of interest: None

Received: 2024-07-06 Accepted: 2024-08-05 Published Online: 2024-08-07
Funding: None

Ethical Approval: Ethical approval for the research protocol was obtained from the Institutional Review Board of the Faculty of Medicine at Marmara University (IRB Approval No: 08.09.2023.1146).

Author Contributions: Conception: FB; SAE; Design: FB; GG - Supervision: FB; GG - Fundings: SAE; Materials: SAE; GG - Data Collection and/or Processing: FB; SAE; GG - Analysis and/or Interpretation: FB; SAE; GG - Literature: FB; SAE; GG - Review: FB; SAE; GG - Writing: FB; SAE; GG - Critical Review: FB; SAE; GG

Main points:
It has been determined that using the smallest possible container and syringe for volume measurements to be performed in a clinical environment using water displacement methods increases the reliability of the results.
ABSTRACT

Objective: This study aims to identify the most reliable method for measuring graft volumes comparable to those harvested from the ramus region using 3D-printed models.

Methods: Using a cross-sectional design in an in vitro setting, CBCT images from 20 individuals who met the inclusion criteria for ramus grafting were examined. Volumetric evaluations were conducted on these images, and 3D-printed graft models were created. Two blinded raters assessed the graft volumes using the displacement method (with beakers of 10 cc, 25 cc, 50 cc capacity and a 100 cc biopsy cup) and the overflow liquid method (with beakers of 10 cc, 25 cc, and 50 cc capacity). The intraclass correlation coefficient and t tests were applied for statistical validation of intra- and inter-rater reliability.

Results: High levels of both intra- and interrater reliability were observed, particularly for the 10 cc rise and overflow methods. These methods exhibited not only exceptionally high ICC values but also statistically meaningful p values. Furthermore, most of these methods strongly correlated and agreed with the CBCT measurements, except for the 50 cc overflow method, which showed significant divergence.

Conclusion: The findings of this study validate the 10 cc beaker methods for reliable 3D printed ramus graft volume measurement and recommend a narrow-diameter syringe for optimal accuracy. These findings have crucial implications for both clinical practice and future research.

Keywords: alveolar ride augmentation, three-dimensional image, organ size
INTRODUCTION
Volume measurement is often crucial, especially in research-oriented studies [1]. The basic application of 3D surface volume measurement in a clinical setting can be traced back to 1967, as outlined by Burke and Beard [2]. Subsequent technological advancements have facilitated the 3D rendering of human tissues through various scanning techniques [3]. Nonetheless, such methodologies carry substantial limitations owing to their complex application procedures [4].

For irregularly shaped objects such as harvested bone grafts, the current method for volume measurement stems from Archimedes' principle, established over two millennia ago. According to this principle, the volume of an object submerged in fluid is equal to the volume of the displaced fluid. Currently, this method serves as the reference test for tracking lymphedema in extremities [5,6], and is also utilized for volumetric calculations postprostatectomy [7], and mastectomy [8], as well as for tumor specimens [9]. In the field of oral and maxillofacial surgery, it is useful in measuring the volume of the temporomandibular joint condyle [10,11], in vitro extracted teeth [12], and bone grafts [13].

While foundational, Archimedes' principle is not without its limitations, and ongoing research continues to explore its intricacies [14]. The principle neglects factors such as surface tension, which can, in practical terms, compromise its premises [15]. Complex liquids can also confound its estimates [16]. In a clinical scenario, the immersion of the graft in a saline solution could influence bone graft volume assessment. Additionally, the method mandates that the diameter of the measurement container be at least as wide as the object, potentially compromising accuracy [14]. Further studies are warranted to determine whether the gold standard 'water displacement method', known for its efficacy in measuring larger limbs, retains its validity and reliability for smaller volumes.

The objective of this study was to identify the most reliable method for the in vitro measurement of graft volume equivalent to that harvested from the autogenous ramus donor site. To achieve this, we used a cross-sectional design to assess the intrarater and interrater reliability of water displacement methods with various containers for measuring the volume of 3D-printed models simulating bone grafts from the ramus region. The ultimate aim is to propose a straightforward, cost-effective method with established validity and reliability for graft volume estimation. The null hypothesis (H0) for this study was that there is no significant difference in the reliability of volume measurements of 3D-printed graft models between the CBCT method and the water displacement methods using various containers.

MATERIAL AND METHODS
Sample Size Calculation
A priori power analysis was conducted using G*Power software (version 3.1) to determine the required sample size for detecting a significant difference between the CBCT and water displacement methods. An effect size (dz) of 0.7 was chosen to represent a medium-to-high effect size, ensuring the detection of a clinically meaningful difference. Assuming an alpha level of 0.05 and a desired power of 0.80, the analysis indicated that a total sample
size of 19 subjects was necessary to achieve adequate statistical power (actual power = 0.822). Therefore, 20 subjects were included in the study.

**Study Design**

The study included 20 consecutive patients who underwent ramus grafting in the Department of Oral and Maxillofacial Surgery Clinic at Marmara University during the year 2022. Inclusion criteria specified that participants must be at least 18 years of age, have a cone beam computed tomography (CBCT) scan, have undergone bone augmentation utilizing the mandibular ramus as the donor site, and provided written consent for the academic use of their radiological data. Patients were excluded if they had systemic diseases impacting bone metabolism, previous mandibular ramus donations, bone-related lesions or surgeries in the ramus area, recent radiotherapy or chemotherapy treatments, or inadequate quality of imaging.

**CBCT Measurements**

Volumetric CBCT measurements were conducted by an observer who was trained and calibrated in CBCT evaluations. CBCT images were obtained from the institutional healthcare system using a uniform machine and settings: Planmeca Promax 3D Mid (Helsinki, Finland), 90 kVp, 10 mA, 10.08 s, 0.20 mm voxel size, and 160x160 mm field of view (FOV). These images were then imported into Slicer 5.2.2 software (Slicer Community) in DICOM format [12,17]. The mandible was segmented based on a threshold value that optimally defined the bone boundaries for each patient [18]. Four osteotomy lines were strategically placed in the mandibular ramus: two vertical lines in the proximal area, one horizontal line apically, and one crestal line, marking the boundaries for ramus graft segments. After osteotomy, the volumes of the graft segments were recorded in cubic millimeters ($\text{mm}^3$). These segments were exported as STL files and imported into Chitubox 1.9.4 software for 3D printing. The 3D models were created with a Phrozen Sonic Mini 4K 3D printer (Phrozen Tech Co., Ltd, Taiwan) using UV-sensitive resin (3D Printing UV Sensitive Resin, Anycubic Technology Co., Ltd, Hongkong). The finished graft models were assigned random numbers for identification.

**Measurement & Evaluation**

One investigator (G.G.) presented the models for assessment. Two blinded raters, unaware of the 3D models they were evaluating and each other's evaluations, carried out two sets of measurements each. The sequence of 3D models examined was randomized using Excel's RAND() function. The primary rater (S.A.E.) conducted an initial measurement and repeated it one week later to calculate intrarater variability; all specimens were reassessed on both occasions. A second rater (F.B.) also executed two separate sets of measurements across the sample to gauge interrater variability.

Volume assessments of the ramus models were performed using the Archimedes principle. Four distinct containers with varying diameters and designs were used: beakers with capacities of 10 cc, 25 cc, 50 cc, and a 100 cc biopsy container. Sterile saline served as the medium for both measurement methods to more closely approximate clinical conditions. Two techniques were used for each container: measuring the volume of displaced liquid and measuring that of overflow liquid. For the biopsy container, only the displacement method was applied. In total, fourteen
unique measurements were taken for each 3D-printed graft model, as each of the two raters carried out measurements twice.

For the overflow group, the container was initially filled to excess, allowing surplus saline to drain until equilibrium was reached. A secondary empty container was placed underneath to collect the overflow upon submersion of the ramus model. The overflow volume was precisely gauged using a 1 cc syringe calibrated to a 0.01 cc scale. A second syringe was used when the ramus model volume exceeded 1 cc. In contrast, the displaced group involved filling the selected containers to a predefined level. The ramus model was then submerged, and the displaced liquid volume was meticulously noted. Consistency in the observer’s eye level was maintained pre- and postimmersion. The volume was also measured using a syringe.

The recorded volumes in cc were converted to mm$^3$ by multiplying by 1000 and compared with values ascertained using Slicer software for additional validation.

**Statistical Analysis**

SPSS (Statistical Package for the Social Sciences) program was used to conduct the statistical analysis (version 22.0, IBM Corp., Armonk, NY, USA). Descriptive statistics were analyzed to summarize the graft volume data obtained from various methods and containers, shedding light on the dataset's central tendencies and spread.

For intrarater reliability, two separate sets of measurements were collected by the first and second raters with a 15-day interval between them. The intraclass correlation coefficient (ICC) was used to measure the level of agreement between these two sets of measurements for each method, providing insight into the repeatability and internal consistency of each rater’s assessments. To assess interrater reliability, both the initial and follow-up measurements from all types of containers were included in the analysis. The degree of consistency and agreement between the first and second raters was evaluated using ICC values, which were categorized according to the classification system outlined by Portney and Watkins: an ICC below 0.50 indicates poor reliability, between 0.51 and 0.75 suggests moderate reliability, between 0.75 and 0.90 indicates good reliability, and above 0.90 signifies excellent reliability.

Scatter plots were created for each pair of measurement methods, starting with the 10 cc container and CBCT, to conduct an initial investigation into the relationships between variables. For each pair, a t test was performed to determine the statistical significance of the differences between the measurement methods. The null hypothesis (H0) assumed that there were no significant differences between the two methods. Bland–Altman plots were used to visually examine the level of agreement between the methods. The mean difference and 95% limits of agreement were calculated. To identify any proportional biases between the pairs of methods, linear regression analyses were conducted. A p value greater than 0.05 led to the retention of the null hypothesis, indicating no proportional bias.
RESULTS

Intrarater Consistency

Table 1 provides an overview of the ICC values, offering insight into the consistency and agreement across the two measurements. The first evaluator's assessments of both the 10 cc rise and 10 cc overflow methods yielded an exceptionally high ICC value of 0.975, supported by p values less than 0.001, signifying strong agreement. Likewise, the first evaluator demonstrated significant agreement in the 25 cc rise and 25 cc overflow methods, with ICC values of 0.746 and 0.864, respectively, each with p values less than 0.001. Evaluation of the 50 cc rise and 50 cc overflow methods also indicated good levels of agreement, manifested by ICC values of 0.876 and 0.580, respectively, each with p values less than 0.02. In contrast, the biopsy container rise method displayed average agreement levels, as indicated by an ICC value of 0.518 and a p value of 0.01 (Table 1).

The second evaluator, assessing the identical set of methods, also observed high ICC values and statistically meaningful p values, denoting strong agreement. For the 10 cc rise and 10 cc overflow methods, the second evaluator reported ICC values of 0.975 and 0.968, respectively, each with p values less than 0.001. The 25 cc rise and 25 cc overflow methods also showed substantial agreement levels, as indicated by ICC values of 0.934 and 0.769, respectively, and p values less than 0.001. The evaluation of the 50 cc rise method yielded a high ICC value of 0.917, with a p value less than 0.001. However, the 50 cc overflow method presented a somewhat lower, yet still noteworthy, ICC value of 0.592 and a p value of 0.001. Finally, the biopsy container rise method, as evaluated by the second evaluator, displayed average agreement with an ICC value of 0.482 and a p value of 0.015 (Table 1).

Table 1. Intrarater reliability between the first (T0) and second measures (15 days later)

<table>
<thead>
<tr>
<th>Method</th>
<th>ICC Coefficient</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First rater</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 cc rise</td>
<td>0.975</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10 cc overflow</td>
<td>0.854</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc rise</td>
<td>0.746</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc overflow</td>
<td>0.864</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc rise</td>
<td>0.876</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc overflow</td>
<td>0.580</td>
<td>0.02</td>
</tr>
<tr>
<td>Biopsy container rise</td>
<td>0.518</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Second rater</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 cc rise</td>
<td>0.975</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10 cc overflow</td>
<td>0.968</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc rise</td>
<td>0.934</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc overflow</td>
<td>0.769</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc rise</td>
<td>0.917</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc overflow</td>
<td>0.592</td>
<td>0.001</td>
</tr>
<tr>
<td>Biopsy container rise</td>
<td>0.482</td>
<td>0.015</td>
</tr>
</tbody>
</table>
**Interrater Consistency**

For the 10 cc rise technique, both the initial and subsequent evaluations indicated high ICC values of 0.961 and 0.897, respectively, each supported by p values less than 0.001. In a similar manner, the 10 cc overflow approach also showed significant agreement across evaluators, with ICC values of 0.897 and 0.952 for the initial and subsequent evaluations, respectively, and p values less than 0.001. For the 25 cc beaker, the ICC values indicated substantial agreement across most metrics. For both the rise and overflow methods, the ICC values for the initial and subsequent evaluations exceeded 0.8, each substantiated by p values less than 0.001 (Table 2).

The 50 cc beaker displayed varying degrees of interrater consistency. The initial evaluation using the rise technique yielded an exceptionally high ICC value of 0.962, substantiated by a p value less than 0.001. The subsequent evaluation revealed an ICC value of 0.936, which was also supported by a p value less than 0.001. In contrast, the overflow method manifested relatively lower ICC values for both evaluations, albeit statistically significant (0.438 and 0.763, with p values of 0.043 and 0.001, respectively). In terms of the biopsy container rise technique, moderate interrater consistency was evident. The initial and subsequent evaluations displayed ICC values of 0.676 and 0.628, respectively, each supported by p values of 0.008 and 0.020 (Table 2).

**Table 2. Interrater reliability between the first and second raters**

<table>
<thead>
<tr>
<th></th>
<th>ICC Coefficient</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First measurement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 cc rise</td>
<td>0.961</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10 cc overflow</td>
<td>0.897</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc rise</td>
<td>0.862</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc overflow</td>
<td>0.898</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc rise</td>
<td>0.962</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc overflow</td>
<td>0.438</td>
<td>0.043</td>
</tr>
<tr>
<td>Biopsy container rise</td>
<td>0.676</td>
<td>0.008</td>
</tr>
<tr>
<td><strong>Second measurement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 cc rise</td>
<td>0.897</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10 cc overflow</td>
<td>0.952</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc rise</td>
<td>0.883</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25 cc overflow</td>
<td>0.816</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc rise</td>
<td>0.936</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50 cc overflow</td>
<td>0.763</td>
<td>0.001</td>
</tr>
<tr>
<td>Biopsy container rise</td>
<td>0.628</td>
<td>0.628</td>
</tr>
</tbody>
</table>
Comparison with CBCT Measurements

The current investigation showed scatter diagrams that revealed significant correlations between various volume measurement techniques—specifically, 10 cc rise, 10 cc overflow, 25 cc rise, 25 cc overflow, and 50 cc rise—and CBCT measurements, excluding the 50 cc overflow method (Figure 1). T tests indicated that the null hypothesis could not be dismissed for all comparisons except the 50 cc overflow versus CBCT, where a p value less than 0.05 led to its rejection. Bland–Altman plots further validated these findings; the majority of data points were within the 95% limits of agreement for all evaluated pairs, except for the 50 cc overflow method. Additionally, linear regression analysis did not reveal any proportional biases across any of the pairs examined, supported by p values greater than 0.05 (Figure 2).

Fig. 1. Scatter plots illustrating the relationship between CBCT and various volume measurement methods.
DISCUSSION
The water displacement method is the reference standard for lymphedema monitoring [5,6], postsurgical volumetric assessments [7], and tumor specimen measurements [9] and has applications in oral and maxillofacial surgery, including temporomandibular joint [10,11], extracted teeth [12], and bone graft volume [13] measurements. This study represents the first comprehensive evaluation of both intrarater and interrater reliability for various volume measurement methods using water displacement in the context of 3D-printed models that simulate bone grafts harvested from the ramus region. We found high degrees of reliability, particularly with the 10 cc beaker methods (both rising and overflow), thus validating their internal consistency and dependability. However, it is crucial to note that as the volume of the container increases, the reliability correspondingly diminishes, indicating potential inaccuracies in larger containers.

The study’s data reveal significant levels of intrarater and interrater reliability across a range of methods employed for measuring the volume of 3D-printed grafts. Intrarater reliability, as illustrated by ICC values that range from moderate to excellent, emphasizes the internal reliability of each measurement approach. Specifically, the 10 cc Beaker (Rising and Overflow) method showed the highest level of intrarater reliability, positioning it as a particularly reliable method for volume measurement. On the other hand, the Biopsy container rise method showed only moderate agreement, indicating inherent limitations or inconsistencies. An observable trend emerged, showing a decrease in reliability as the size of the container increased, which calls for further investigation into potential errors in larger containers.

Patterns in interrater reliability also became evident across varying container sizes and measurement techniques. Exceptionally high ICC values were observed for both the rising and overflow methods in the 10 cc container,
confirming the stability of these methods. This consistency is particularly remarkable, with ICC values exceeding 0.9 for most samples, highlighting significant agreement between evaluators. The results for the 25 cc cup supported the reliability of both the rising and overflow techniques. However, the 50 cc container revealed different outcomes: while the rising method maintained a high level of agreement, the overflow method evidenced lower levels of interrater reliability. This observation aligns with our earlier finding that reliability tends to decline with larger containers. Consistent with intrarater findings, the Biopsy container method showed only moderate levels of interrater reliability, suggesting caution is advised until further validation. The high level of interrater reliability, especially with smaller container sizes, underscores the suitability of these methods for scenarios requiring precise data collection.

Comparisons with cone-beam computed tomography (CBCT) generally demonstrated good agreement for several methods, such as the 10 cc and 25 cc rising methods, which were also substantiated by high intrarater and interrater reliability figures. However, the 50 cc overflow method did not align well with CBCT and displayed lower reliability, indicating that this method may require further evaluation.

Limitations
One limiting factor for the use of the water displacement method is surface tension. In a clinical setting, the surface tension increases when saline is used instead of water. While the addition of ethanol to water has been suggested to reduce this surface tension, such a practice is not feasible for calculating the volume of autogenous ramus in clinical applications [19]. Another consideration is related to the materials used in this study. The measurements were conducted on 3D-printed models made of acrylic, which may not perfectly simulate the characteristics of autogenous ramus grafts, thus necessitating caution when interpreting these findings to a clinical context. Additionally, the markings on the side of the container are also of importance; a wider range can decrease the reliability of measurements. As the needed area of the container expands—necessary to accommodate the ramus model—a need for containers with larger levels may arise. However, an alternative approach involving the use of a narrow-diameter syringe to draw fluid has demonstrated a positive effect on result reliability.

CONCLUSIONS
In summary, this study supports the reliability and validity of using a 10 cc beaker with water displacement methods for assessing 3D printed autogenous ramus graft model volumes from the mandibular ramus. We recommend the use of the smallest possible container that can accommodate the graft and the use of a narrow-diameter syringe for extracting the displaced liquid to ensure the most accurate and reliable measurements.
REFERENCES


