MIRTHE VAN DER KOOIJ

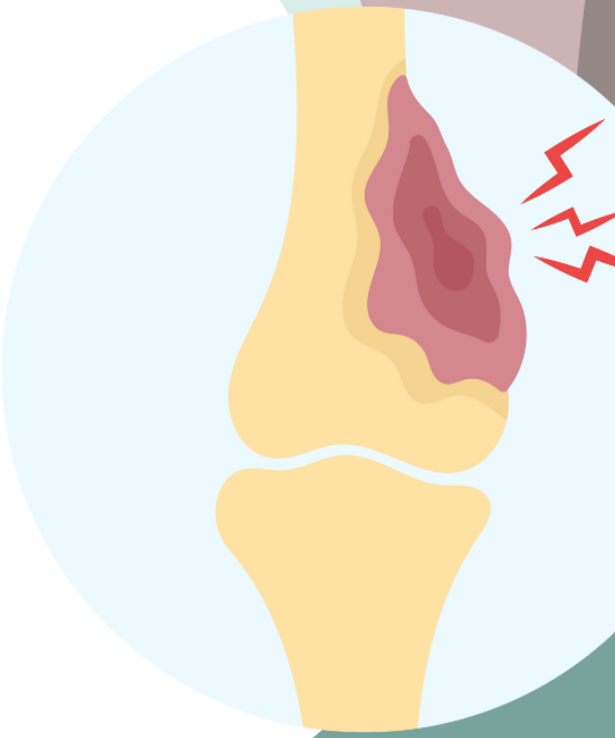
Achieved accuracy of 3D surgical plans in complex orthopedic procedures

Thesis Report

A thesis submitted to
Delft University of Technology,
Erasmus University Rotterdam
and Leiden University

In partial fulfillment of the
requirements for the degree of

MSc. Technical Medicine
Track Imaging & Intervention



3 JULY 2025

[This page intentionally left blank]

Achieved accuracy of 3D surgical plans in complex orthopedic procedures using computer assisted surgery or patient specific instruments

MSc thesis Technical Medicine

By

Mirthe van der Kooij
Student number : 4652258
3th July, 2025

Thesis in partial fulfilment of the requirements for the joint degree of Master of Science in

Technical Medicine

Leiden University ; Delft University of Technology ; Erasmus University Rotterdam

Master thesis project (TM30004 ; 35 ECTS)

Dept. of Biomechanical Engineering, TUDELFT

04-11-2024 – 03-07-2025

Supervisors:

Technical supervisor:	Dr. Ir. B.L. Kaptein	Biomechanical engineer
Medical supervisor:	Drs. D. Broekhuis	Orthopedic surgeon
Medical supervisor:	Dr. J.G. Gerbers	Orthopedic surgeon

Thesis committee members:

Dr. Ir. B.L. Kaptein	LUMC	Chair & Technical Member
Dr. J.G. Gerbers	LUMC	Medical member
Drs. D. Broekhuis	LUMC	Medical member
Dr. J.W. Colaris	Erasmus MC	External member

Universiteit Leiden The Netherlands

TU Delft Delft University of Technology

Erasmus University Rotterdam

[This page intentionally left blank]

Preface

Eight months ago I started my Master's thesis, and this document marks its successful completion and the end of my time as a student!

When I began my Bachelor's degree at TU Eindhoven, I developed a keen interest in orthopaedics. However, the physical patient component was lacking in this program, which led me to switch to Technical Medicine. After completing a pre-master's program, I finished my studies three years later. I really enjoyed my Master's program, particularly the two years of clinical research internships where I experienced firsthand the added value of a Clinical Technologist across multiple hospitals.

During my Master's thesis research, I was eager to combine my interests in orthopaedics and oncology with image processing and programming. This field represents a specialized niche, so I consider myself very lucky to have had this opportunity at the LUMC. I genuinely enjoyed my time there, where I had the privilege to work within the technical field of orthopaedics alongside a complex patient population and a welcoming multidisciplinary team. This experience allowed me to develop both clinically and technically. I believe we can be very proud of what we achieved as a team, and I hope our work will contribute to evaluating the quality of complex orthopaedic surgical procedures.

I am very grateful for my three supervisors: Bart Kaptein, Demien Broekhuis and Jasper Gerbers. Their guidance and support throughout my thesis were invaluable. Bart provided a technical perspective on my work, demonstrating patience and analytical sharpness while offering critical technical insights that helped me develop my technical skills. Demien and Jasper provided essential clinical perspectives on my research. A special thanks to Jasper for his endless enthusiasm and remarkable patience.

Finally, I would like to thank my partner, Johan Hennipman, who consistently supported me throughout my studies.

Now that my time as a student has come to an end, I am grateful for all the amazing opportunities I have had, and I eagerly look forward to the exciting opportunities that lie ahead!

Mirthe van der Kooij
Haarlem, July 2025

Summary

Preoperative planning has become standard of care in complex orthopedic procedures. Procedural accuracy is crucial for effectively executing the plan. However, quantitatively evaluating whether the preoperative plan is successfully performed remains a challenge due to the lack of a standardized measure. Many studies have examined the deviations between the 3D planned and achieved planes using surgical saw blades. However, the various methods used to quantify these deviations make it difficult to compare results across studies.

First, a scoping review was conducted to systematically examine the various methods used to define achieved planes and quantify linear and angular deviations between the planned and achieved planes. This review identified promising methods for standardization. Second, a research protocol was written to obtain ethical approval from the ethics committee to execute our study. Last, the identified promising methods were applied to our dataset, and the procedural accuracy of the resected bone was evaluated relative to the preoperative planning.

To our knowledge, this is the first study to compare different methods for assessing procedural accuracy in orthopedic surgery. Our findings indicate that the point cloud method appears to be a good alternative to the mesh method for identifying the achieved plane. Furthermore, the average outcome measures observed in this study provide a useful baseline for assessing procedural accuracy in complex orthopaedic procedures using 3D planning and computer assisted surgery (CAS).

Table of Contents

List of Abbreviations

Abbreviation	Meaning
SM	Surgical margin
ESM	Error surgical margin
L	Location accuracy
F	Flatness
3D	Three-dimensional
CT	Computed tomography
MRI	Magnetic resonance imaging
PSI	Patient-specific instrument
CAS	Computer-assisted surgery
SLS	Selective laser Sintering
COG	Center of gravity
2D	Two-dimensional
TKA	Total knee arthroplasty
ICP	Iterative Closest Point
SVD	Singular value decomposition
mm	Millimeters
PCA	Principal component analysis
CI	Confidence interval

Definitions

Osteotomy	Planned/performed cutting plane in bone (Fig. 1)
Achieved plane	Cutting plane achieved after surgical resection (Fig. 1)
Planned plane	Cutting plane planned preoperatively using 3D imaging (Fig. 1)
Deviations within achieved plane	Local variations within the achieved plane. Only applicable if the cutting plane is uneven
Linear deviation	The distance (in mm) between the achieved and planned planes
Angular deviation	The angle (in degrees) between the achieved and planned planes
Surgical margin (SM)	The minimum distance (mm) between the planned or achieved plane and the boundary of the tumor (Fig. 1A)
Error surgical margin (ESM)	The difference (mm) between SM and the desired safe margin: negative for closer to tumor than planned (R1,R2), positive for a conservative resection (R0), and zero for matching exactly the planned plane
Kerf	The loss of bone material that arises during an osteotomy (Fig 1B.)

<u>Parameters in accordance with the ISO 1101 standard - Geometrical product specifications (GPS)</u>

Location accuracy (L)	The maximum distance (mm) between the achieved and the planned plane (Fig. 1A)
Flatness parameter (F)	The minimum distance (in mm) between two parallel planes that include the achieved plane (Fig 2.)

<u>Parameters in accordance with the ISO 5725 standard - Accuracy (trueness and precision) of measurement methods and results)</u>

Precision	The degree of variability in the resection cuts
Accuracy	The linear and angular deviation of the achieved plane from the preoperatively planned plane.

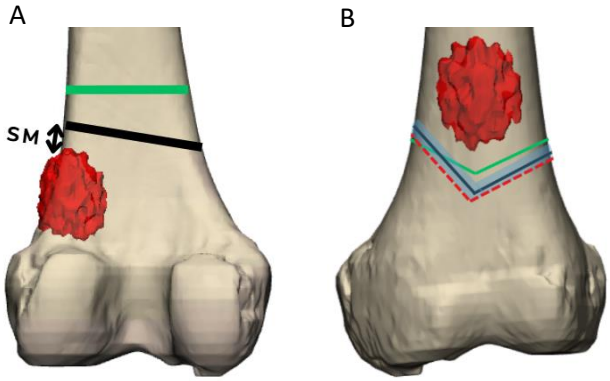


Figure 1: Resection osteotomy. Black indicates the achieved plane. Green represents the planned plane. (A) Location (L) and surgical margin (SM) parameters are used to evaluate the achieved plane. (B) Definition of kerf. The thick blue line shows the kerf, which is the bone loss that arises during an osteotomy. Red dotted line is the measured plane.

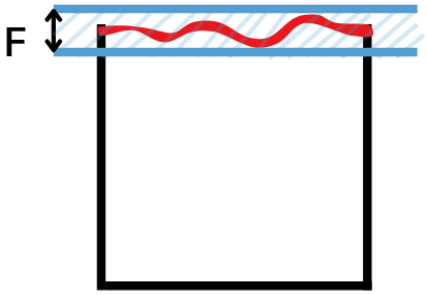


Figure 2: Red indicates the achieved plane. Blue represents the two parallel planes. Flatness (F) is defined as the minimum distance (mm) between the two planes.

[This page intentionally left blank]

I

TM Literature Study

Outcome measures to quantify the differences between planned and achieved osteotomy planes in computer-assisted surgery: a scoping review

Abstract

BACKGROUND: Preoperative planning has become standard of care in complex orthopedic procedures, demonstrating significant value in joint preserving resections and complex reconstruction cases. This planning process requires three-dimensional (3D) imaging, usually computed tomography (CT), sometimes combined with magnetic resonance imaging (MRI), to create a 3D bone model. This enables the surgeon to visualize the anatomy in 3D and virtually plan the procedure. The 3D model can then be used preoperatively by creating patient-specific instruments (PSIs) or using computer-assisted surgery (CAS). Procedural accuracy is crucial for effectively executing the plan. However, quantitatively evaluating whether the preoperative plan is successfully performed remains a challenge due to the lack of a standardized measure. Many studies have examined the deviations between the 3D planned planes and the achieved planes achieved using surgical saw blades. However, the various methods used to quantify these deviations make it difficult to compare results across studies.

OBJECTIVE: The present study undertakes a scoping review of research on the determination of linear and/or angular deviations between the achieved and planned planes in bone (tumor) resections, aiming to establish a more standardized approach for evaluating procedural accuracy.

DESIGN: Online databases were searched to identify papers published up to 2024, resulting in the selection of 36 publications that used quantitative outcome measures.

RESULTS: The majority of publications reported distance (error) as linear outcome measure (63%) and angle (error) as angular outcome measure (75%). In studies with an oncological focus, the majority reported distance (error) as linear outcome measure (63%) as well, while roll-pitch-yaw angles and angle (error) are used equally as angular outcome measures. Distance (error) outcome measures, including the mean distance between points, and the mean and median distance between center of gravity (COG), as well as cutting accuracy outcome measures, including surgical margin (errors) and flatness have the most potential based on established parameters. Mean angle (error) outcome measures, including angle between normal vectors of the achieved and planned plane, and roll, pitch and yaw angles have the most potential based on the same parameters. The most promising linear and angular outcome measures in oncological studies include the mean distance between achieved and planned planes using points, surgical margin (error), and roll-pitch-yaw angles.

CONCLUSIONS: There is a lack of standardized methods and outcome measures in research on cutting planes. This study presents a proposal for standardization. Future research should determine which method and outcome measures are the most suitable for standardization.

KEYWORDS: bone resection, computer-assisted surgery, linear deviation, angular deviation, cutting plane, surgical accuracy, systematic

1.0 Introduction

Conventional preoperative planning methods use two-dimensional (2D) images obtained from radiography, CT, MRI or ultrasound to plan the cutting plane (1). Advances in CT imaging now allows for three-dimensional (3D) image reconstruction through volume rendering (1), providing a more comprehensive understanding of the patient's anatomy. Despite these advancements, the gold standard in presurgical planning remains the visualization of 2D images. However, in complex orthopedic procedures that require higher precision, such as multiplanar osteotomies and complex tumor resections, 3D reconstructed bone models are used (2-4). These models are derived from computed tomography (CT), sometimes combined with magnetic resonance imaging (MRI) and enables the surgeon to visualize the anatomy in 3D and virtually plan the procedure (1-4). This 3D model can then be used preoperatively to develop a patient specific instrument (PSI) or a computer-assisted surgery (CAS) plan for navigation purposes. Similar practices are in use in orthognathic surgery planning (3). Several techniques are described to perform these plannings, including navigation systems and PSIs (5, 6).

In the early 1990s the first-generation navigation systems were introduced, enabling real-time tracking of instruments using preoperative imaging (3, 4). Initially used in neurosurgery, these systems are now in standard practice in various surgical specialties, including otolaryngology, craniomaxillofacial surgery, and orthopedics (4). PSIs were developed as an alternative to navigation systems. The concept began in the early 2000s with total knee arthroplasty (TKA) (5). Nowadays, PSIs are also used in bone tumor resections, correction osteotomies and other orthopedic procedures (6).

Similar results have been reported for surgical navigation and PSIs (7). Both technologies have advantages compared to conventional methods, including reduced surgical time, exposure and blood loss, as well as improved accuracy, resulting in better surgical margins and patient outcomes (7-9).

However, each technology has advantages and disadvantages which are summarized in Table 1. One advantage of navigation systems include their availability on short notice (7). In contrast, PSIs require significant time and resources for preoperative planning and the fabrication of guides PSIs, on the other hand, are vulnerable to intraoperative errors, due to inadequate bone surface preparation and guide placement errors, which cannot be detected in real-time unless navigation is used alongside PSIs (7, 10). Furthermore, navigation systems typically require less soft tissue dissection (10). A disadvantage of PSIs is the need for anatomical landmarks to ensure proper positioning, which is crucial for the accuracy of the surgical steps (11). This can be particularly challenging in cases such as diaphyseal localization. In these cases, adequate support from the diaphysis to the metaphysis is required to prevent positioning errors along the length axis.

In contrast to navigation, PSIs have the advantage to guide the surgical saw blade along a predetermined resection path (10). Therefore, PSIs may be superior in resection precision to

navigation in bone tumor resections (10). This is particularly beneficial for geometric and multiplanar osteotomies, as more host bone is preserved for reconstruction, and a better limb function has been reported (10, 12). Other advantages of PSIs compared to navigation systems include reduced intraoperative time (7, 13), lower radiation exposure and the fact that there is no need to invest in an expensive navigation system. PSIs are more cost-effective, as it is an accessible technology with a minimal learning curve (7, 13).

Quantitatively evaluating whether the preoperative plan is successfully performed remains a challenge due to the lack of standardized measurements. Many studies have investigated the deviations between the 3D planned planes and the achieved planes achieved (2, 14-48). However, there are various methods used to quantify these deviations, making it more difficult to compare studies with each other. Therefore, the purpose of this review is to summarize the methodologies used to describe the achieved plane, and the associated linear and angular outcome measures, in order to investigate the best suitable approach for evaluating the procedural accuracy of resected bone compared to the preoperative plan.

In this review, studies with an oncological focus are discussed separately, as the postoperative accuracy of the preoperative plan has a greater clinical impact compared to studies on correction osteotomies, which may result in different methodologies and outcome measures.

For simplicity, CAS and/or PSI will be used as the overarching terms to refer to computer-generated 3D modeling, navigation, and PSIs in this review.

Table 1: Advantages and disadvantages of navigation systems and patient-specific instruments.

	Navigation systems	**Patient-specific instruments**
	Available on short notice	Increased preoperative time and resources
	Ability to make intraoperative adjustments	Limited ability to make intraoperative adjustments
	Less soft tissue dissection, resulting in smaller incisions	Intraoperative errors including bone surface preparation and guide placement errors
	Increased surgical time	Reduced surgical time
	Steep learning curve	Minimal learning curve
	Costs	Accessible technology
	More radiation exposure	Lower radiation exposure

2.0 Materials and Methods

A systematic literature search was performed on November 20, 2024, following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (49). Six databases were searched for relevant studies: PubMed, EMBASE, Web of Science, Cochrane library, Emcare and Academic Search Premier. The search strategies were drafted by an experienced librarian J.W.S. and further refined through team discussion. The final search strategy for the PubMed database can be found in Appendix A. The final search results were exported into EndNote, and duplicates were removed by J.W.S.

Studies describing correction osteotomies or bone tumor resections were both included and initially considered together. The inclusion criteria were studies that evaluated the difference between the achieved and planned plane using CAS in bone, and they needed to report quantitative outcome

measures. Only studies published in English or Dutch were considered. Articles determining total reconstruction or navigation accuracy were excluded. Other exclusion criteria included the use of modalities other than CT and MRI, and grey literature. The in- and exclusion criteria can be found in Table 2. One reviewer (M.K) reviewed all articles found in the search based on title and abstract. A second review of full-text articles was performed when the abstract was unclear.

The included studies were screened on their study characteristics. Furthermore, the methodology to determine the achieved plane was extracted, as well as the description of deviations within the achieved plane and the deviations between the achieved and planned plane (e.g. linear and angular outcome measures). For 3D methods, registration between preoperative plan and postoperative result is required before comparisons between the achieved and planned planes can be made and quantified. However, the specific registration methods were not described in this study. Subsequently, a subjective scoring system for evaluating papers based on grounding criteria was developed to compare different methodologies using key parameters. These parameters, identified and scored in consultation with an orthopedic surgeon and a biomechanical engineer, include complexity, implementation, accuracy, precision, and local variations. Complexity refers to the ease of understanding of the methodology for non-technicians. Implementation addresses the feasibility of integrating the methodology from both clinical and theoretical perspectives. Accuracy and precision are defined in accordance with the ISO 5725 standard (50), and local variations refers to the ability to describe complex, uneven planes.

No quality assessment was performed, as our interest was on the outcome measures describing differences between achieved and planned planes, making the quality of the article less significant.

Table 2: In- and exclusion criteria for the scoping review

Inclusion criteria	**Exclusion criteria**
- Quantification of cutting planes	- Total reconstruction accuracy
- Skeletal system	- Navigation accuracy
- Available in English or Dutch	- Imaging of other modalities than CT or MRI
- Quantitative outcome measures	- Grey literature

3.0 Results

The systematic search initially identified 339 records, of which 186 were duplicates. Additionally, one study was identified through snowballing, resulting in a total of 154 records. After screening titles and abstracts, 50 studies were excluded. Following a full-text review, 68 articles were excluded, leaving 36 studies eligible for inclusion in this review. Among these, 19 studies have an oncological focus, see Figure 3.

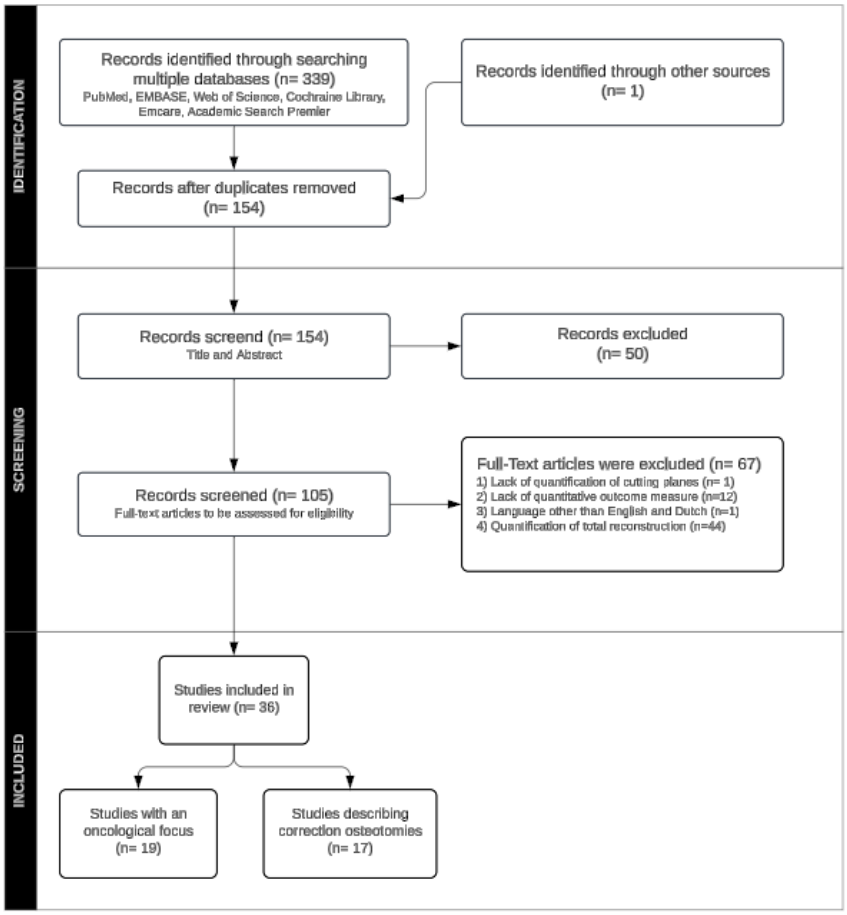


Figure 3: Overview of systematic search, resulting in 36 included articles.

3.1 Study characteristics

Study characteristics include type of study (clinical, sawbones, cadaver or animal), scanned subject (patient or resection specimen), and type of pre- and postoperative scan.

Seventeen studies are clinical (2, 15, 16, 21, 23-27, 29, 31, 37, 38, 43, 44, 46, 47), twelve are sawbones (18, 19, 28, 32-36, 39, 42, 45, 48), five are cadaveric (14, 17, 22, 30, 40) and two are animal studies (20, 41). Nine studies scan resection specimens postoperative (2, 20, 30, 31, 38, 45-48), while the others scan patients.

Six studies use combined CT and MRI scans (2, 25, 27, 31, 38, 46), one study uses a preoperative high-resolution laser scan (28), and the remaining studies use only preoperative CT scans. Thirty-one studies use a postoperative CT scan (14-18, 21-27, 29-45, 47, 48), one uses a postoperative high resolution laser scan (28), one study uses a postoperative MRI scan (2), one study does not describe the imaging method (19), and the remaining two studies do not use a postoperative scan, because they determine the outcome measures intraoperatively from the resection specimen (20, 46). For an overview see Table 3. For a complete overview, refer to Table 12 in Appendix B and Table 13 in Appendix C. The latter specifically outlines the study characteristics for studies with an oncological focus.

Table 3: Overview of the study characteristics

First author and year	Type of study	Scanned subject	Preoperative scan	Postoperative scan
Qu, 2015	clinical	patient	CT	CT
Otsuki, 2013	clinical	patient	CT	CT
Brouwer de Koning, 2021	clinical	patient	CT	CT
De Stavola, 2022	clinical	patient	CT	CT
Vanesa, 2021	clinical	patient	CT	CT
Weijs, 2016	clinical	patient	CT	CT
Fujiwara, 2018	clinical	patient	CT	CT
Hasan, 2020	clinical	patient	CT	CT
Hoving, 2018	clinical	patient	CT + MRI	CT
Liu, 2024	clinical	patient	CT	CT
Gouin, 2014	clinical	patient	CT + MRI	CT
Bai, 2014	clinical	patient	CT	CT
Gerbers, 2013	clinical	patient	CT	CT
Evrard, 2022	clinical	resection specimen	CT + MRI	MRI
Müller, 2020	clinical	resection specimen	CT + MRI	CT
Wong, 2013	clinical	resection specimen	CT + MRI	N/A
Ritacco 2013	clinical	resection specimen	CT + MRI	CT
Wong, 2015	clinical	resection specimen	CT	CT
Pietruski, 2019	sawbones	patient	CT	CT
Pietruski, 2019	sawbones	patient	CT	CT
Pietruski, 2020	sawbones	patient	CT	CT
Pietruski, 2023	sawbones	patient	CT	CT
Ter Braak, 2020	sawbones	patient	CT	CT
Winnand, 2022	sawbones	resection specimen	CT	CT
Chan, 2022	sawbones	patient	CT	unknown
Khan, 2013	sawbones	patient	high-resolution laser scanner	high-resolution laser scanner
Cartiaux, 2014	sawbones	patient	CT	CT
Siegel, 2020	sawbones	patient	CT	CT
Zhao, 2024	sawbones	resection specimen	CT	CT
Ackermann, 2021	cadaveric	patient	CT	CT
Dobbe, 2014	cadaveric	patient	CT	CT
Caiti, 2021	cadaveric	patient	CT	CT
Modabber, 2022	cadaveric	resection specimen	CT	CT
Sternheim, 2015	cadaveric	patient	CT	CT
Cho, 2018	animal	resection specimen	CT	N/A
Sun, 2021	animal	patient	CT	CT

3.2 Determination of achieved plane

There are several 2D and 3D methods to determine an achieved plane. In 3D, a plane is defined by a point and a normal vector. A simple approach involves manually positioning a plane within a 3D model.

However, a more accurate approach involves positioning at least three points on the plane and subsequently fitting a plane through those points. Positioning points on a plane can be achieved through several techniques as outlined in the 3D method subsection below. Of the 36 included studies, 19 studies describe their method for determining the achieved plane (14, 16-18, 20, 22, 23, 25, 26, 28-30, 38, 40, 42-46). Two studies use 2D methods (20, 46) while the remaining studies use 3D methods.

3.2.1 2D method

Two studies determine the achieved plane manually, where it is visible intraoperatively (Fig. 2) (20, 46). The study by Wong et al. (2013) assessed the achieved plane by comparing the cross-section of the resected bone with complementary prosthesis templates (Fig 4A.) (46). This approach focused exclusively on the size and shape of the plane. A deviation exceeding 2 mm in any dimension was considered inaccurate. The study by Cho et al. (2018) evaluated the achieved plane of two sections of the resection specimen (Fig. 4B) (20).

Advantages of the intraoperative 2D methods include direct insight into the quality of the resection specimen and/or surgical margins. Additionally, no postoperative CT scan is needed, resulting in no additional radiation exposure to the patient. However, precision and accuracy are potentially lower compared to 3D methods and 2D methods are likely less effective at describing complex, multiplanar planes. Furthermore, the description of the achieved plane is less accurate when using the method described by Wong et al. (2013) (46). For an overview of the comparisons between the 2D and 3D methods, see Table 4. For an overview of the methods used in each study, see Table 5.

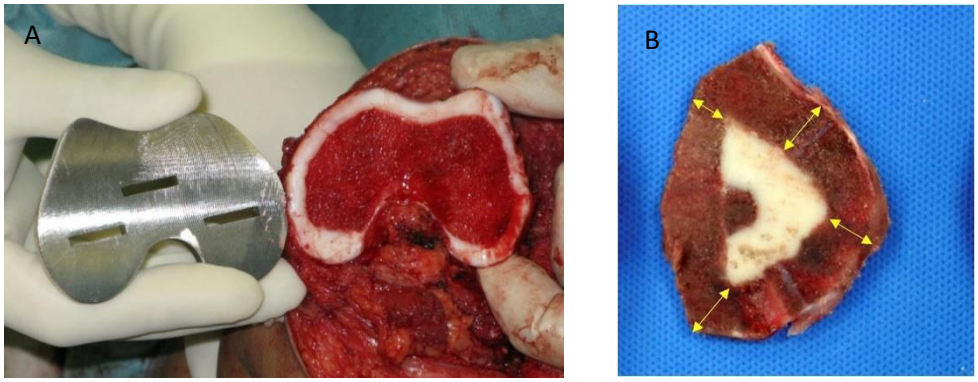


Figure 4. Intraoperative determination of achieved plane. (A) Prosthesis template used to measure bone resection accuracy near the joints. (B) Resection error measured at four locations. Figures copied from the study by Wong et al. (2013) (46) (A) and Cho et al. (2018)

3.2.2 3D method

17 studies determine the achieved plane in 3D, and are categorized into three methods; digitizing, manual 3D plane placement, and selecting points followed by a plane fit.

One study belongs to the digitizing method, in which achieved planes are determined by digitizing them on the resected bone model directly, using a coordinate measuring machine (18). This is the most accurate method, with the highest precision. Additionally, this method does not expose the patient to additional radiation and is capable of describing local variations within the achieved plane. However, it is used in an experimental setting, making it more difficult to implement in a clinical environment. Additionally, this method requires advanced and expensive equipment, and it might be more complex to understand for non-technicians. The manual plane placement method consists of two studies (23, 25). The study by Fujiwara et al. (2018) determined the achieved plane using multiple slices in a postoperative CT scan(23) (Fig. 5), Fig. 5), while the study by Gouin et al. (2014) determined

the achieved plane by placing a 3D plane within a 3D model using a 3D medical viewer (25). The manual plane placement method is relatively simple to implement in a clinical environment, but it may be less accurate than the digitizing method. Additionally, there are no clear specifications for the placement of the plane, making the precision of this method lower. This also results in a more complex method to understand. Moreover, this method is not effective in describing local variations within the achieved plane.

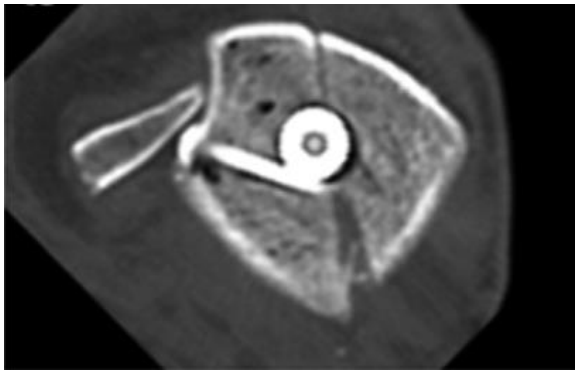


Figure 5. Determine achieved plane from postoperative CT scan. Figure copied from the study by Fujiwara et al. (2018) (23).

The remaining studies describe the achieved plane using the point selection method. Points are selected manually (14, 16, 17, 22, 25, 26, 29, 30, 40, 43-45) or randomly (28, 38, 42) selecting points in a 3D model or axial slices from a postoperative CT scan (26, 40, 44). Subsequently, a plane is fitted through those points. These methods mainly differ in the number of points selected and the approach used, see Table 4. The point selection method provides clearer specifications for placing the points, making it more precise than the manual plane placement method. Furthermore, precision improves with the number of points used, allowing it to effectively describe local variations within the achieved planes. This method does not require a 3D coordinate measuring machine, making it more simple to implement in a clinical environment. Additionally, it is less complex to understand for non-technicians.

Among the studies that manually select points, two select a single point on the achieved plane in a 3D model (16, 22). In the first study, they use a semi-automatic approach by determining the outward-facing normal vector after selecting the point (16). All neighboring points with normal vectors deviating less than 10 degrees are considered to belong to the achieved plane (Fig. 6). This is the most easy method to implement in clinical and theoretical perspectives, as it is semi-automated. In the second study, they use a smoothing approach by manually positioning one or multiple 3D spheres with radii ranging from 2 to 12 mm, increasing in 2 mm increments, on regions of the achieved plane (22). By adjusting the radius, a larger or smaller part of the achieved plane can be selected. The enclosed surface patches are then used to determine the parameters of the best-fitting plane. The points within each surface patch represent the points of the achieved plane. To account for local variations within the achieved plane, multiple 3D spheres and their corresponding patches can be used (Fig. 7). This method does not rely on the choice of the initial starting point, making it more precise. Additionally, the adjustable size of the 3D sphere offers better control over the number and distribution of the selected points. However, this subgroup is unable to describe the contour of the bone, which makes it less accurate. Additionally, this subgroup is time-consuming, making it more challenging to implement from theoretical perspectives.

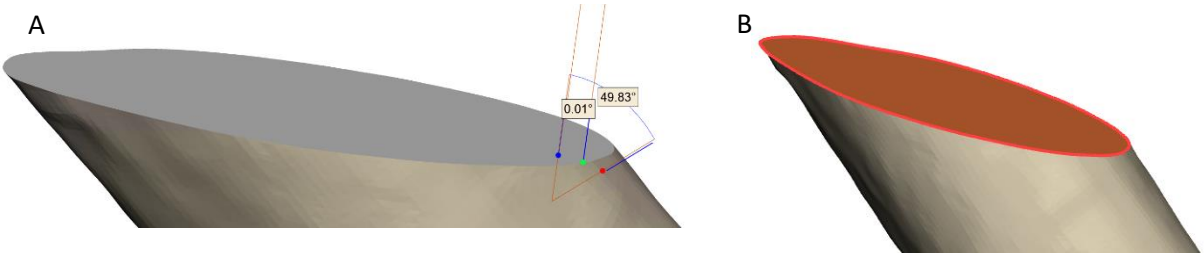


Figure 6: Normal vector method. (A) The grey platform represents the achieved plane. Blue point indicates the manually selected single point. Green point lies within the achieved plane (<10°), while the red point lies outside the achieved plane (>10°). (B) Achieved plane (orange) is automatically fitted through all green points.

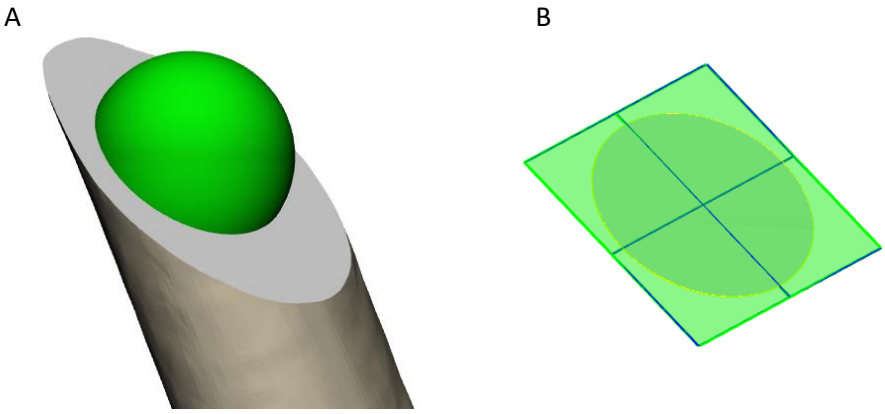


Figure 7: 3D sphere point method. (A) 3D sphere of adjustable size (green) is positioned on the achieved plane of the femur. (B) Achieved plane is fitted through a surface patch of the 3D sphere.

Among the studies that manually select points, nine select multiple points (14, 17, 26, 29, 30, 40, 43-45). In one study they select an unknown number of points on the achieved plane in the postoperative CT scan (26). Three studies select three points on the achieved plane in a 3D model (17, 29, 43) (Fig. 8). The three-point method is easy to understand for non-technicians. However, selecting more points results in a more accurate description of the achieved plane. Two studies select four points (14, 40). The study by Ackermann et al. (2021) selects four landmarks points in a 3D model (14), while the study by Sternheim et al. (2015) selects two entry and two exit points in a postoperative CT scan (Fig. 9) (40). Another study selects 25 landmark points on the achieved plane in the postoperative CT scan (44). These methods provide clearer specifications for placing the points, making it more precise than the three-point method. The final two studies select manually mesh portions on the achieved plane in a 3D model (30, 45) (Fig. 10). The mesh portions method divides the achieved plane into small sections

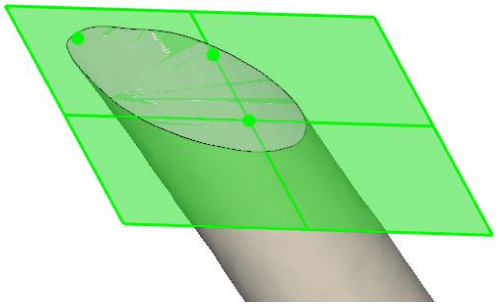


Figure 8: Three points are manually selected on the achieved plane, followed by fitting a plane through these points.

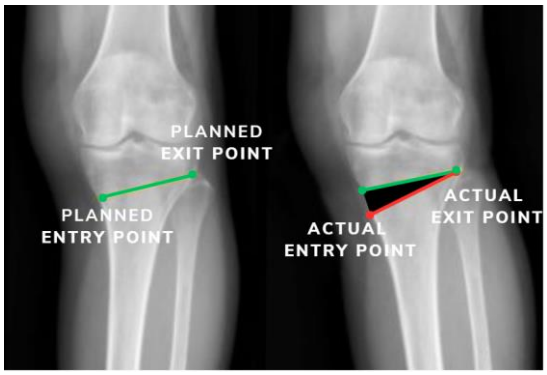


Figure 9. Selecting 4 points on the postoperative CT scan to visualize the planned plane and to determine the achieved plane.

and describes each section separately, making it more accurate. Additionally, it selects more points, allowing for a better representation of local variations within the achieved plane. However, this method is more complex to understand for non-technicians.

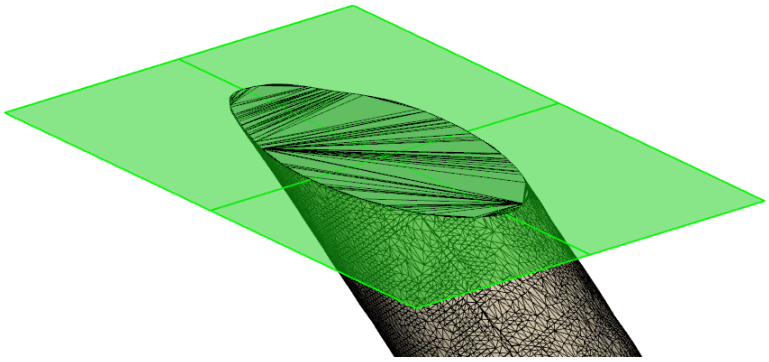


Figure 10: Triangular structures are visualized on the achieved plane, followed by fitting a plane through the vertices of these triangles.

The studies that randomly select points on the achieved plane use different methods. All three studies select multiple points in their 3D models. The first study determines the outward-oriented normal vectors from the achieved plane and intersect them with the planned plane (42). All normal vectors within a distance approximately equal to the thickness of the bone resection are considered to belong to the achieved plane (Fig. 11). This results in points on both sides of the achieved plane. This method is not automated, resulting in a lower precision. The second study converts the 3D model into closely spaced, discrete points on the outer cortices of the achieved plane using reverse engineering software (28) (Fig. 12). The discrete point method does not select points within the achieved plane, making it incapable of describing local variations within this plane. The third study converts the 3D model into a point cloud, resulting in points on the achieved plane (Fig. 13) (38). The discrete point and point cloud methods are the most accurate, as they divide the achieved plane into smaller sections.

For an overview of the comparisons between the 2D and 3D methods, see Table 4. For an overview of the methods used in each study, see Table 5. For a complete overview, refer to Table 12 in Appendix B.

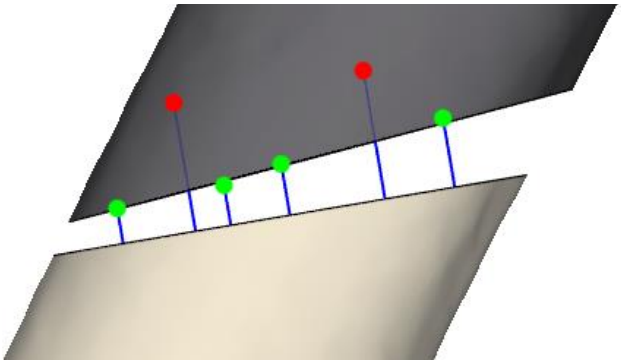


Figure 11: Grey represents the planned proximal part of the femur, while beige represents the achieved distal part of the femur. Red points indicate regions where the normal vector (blue) exceeds the thickness of the bone resection, whereas green points fall within the acceptable range.

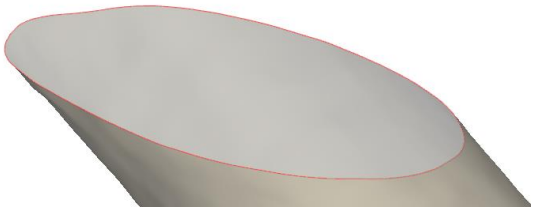


Figure 12: Multiple points on the outer cortex of the femur are selected.

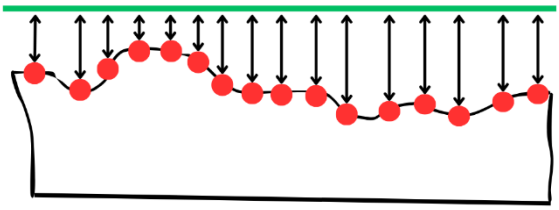


Figure 13. 3D model converted into point cloud model. Green indicates planned plane. Red indicates points on achieved plane. Arrows indicate distance between planned and achieved plane.

Table 4: Subjective scoring of papers based on grounders to compare 2D and 3D methods for describing the achieved plane. (-) indicates poor performance, (+/-) indicates average performance, (+), indicates good performance, and (++) indicates excellent performance (e.g. '+' for complexity means the method is not complex, and '+' for implementation means the method is easy to implement).

	2D	**3D**									
	M1	**M1**	**M2**	**M3A – manual**		**M3B - manual**			**M3C - random**		
	Intraoperative	Digitalizing	Manual determination	Normal vector	3D sphere	3 points	Multiple points	Mesh portions	Normal vector comparison	Discrete point conversion	Point cloud
Complexity	+	-	+/-	+	+	++	++	-	-	+	+
Implementation	+	--	+	++	+/-	++	+	++	+/-	+/-	+/-
Accuracy	-	++	+/-	+	+/-	+/-	+	++	+	++	++
Precision	-	++	-	+/-	+	-	+/-	++	+	++	++
Local variations	-	++	-	+/-	+	-	+	++	+	--	++

Table 5: Overview of the methods used to determine the achieved plane per study

Achieved plane	**First author and year**
Determined in 2D	
- **Visible intraoperative**	Wong, 2013; Cho, 2018
Determined in 3D (alternative)	
- **Coordinate measuring machine**	Cartiaux, 2014
- **Manual**	Gouin, 2014; Fujiwara, 2018
Determined in 3D	
- **Selecting points & plane fit**	
○ **Single point**	Brouwer de Koning, 2021; Dobbe, 2014
○ **3 points**	Liu, 2024; Vanesa, 2021; Caiti, 2021
○ **4 points**	Ackermann, 2021; Sternheim, 2015
○ **Mesh portions**	Modabber, 2022; Winnand, 2022
○ **Multiple points (manual)**	Weijs, 2016
○ **Multiple points (random)**	Khan, 2013; Ter Braak, 2020; Ritacco, 2013 Hasan, 2020
○ **Unknown amount**	

3.3 Deviations within achieved plane

A cutting plane is not necessarily flat. There are several methods to determine the local variations in the achieved plane. Twenty-four of the 36 studies do not describe local variations within the achieved planes. The remaining 12 studies are divided into four methods; intraoperative navigation system, flatness parameter, dividing into several planes, and selecting multiple points.

The intraoperative navigation method is used in one study (46). Wong et al. (2013) determined the local variations within the achieved plane using an intraoperative navigation system by positioning virtual screws along the planned plane (46). This method cannot detect local variations when the navigation system is used incorrectly or is poorly positioned. Two studies use the flatness parameter method to evaluate the position of the achieved plane with respect to the planned plane, see Figure 2 (18, 28). This method is easy to implement from both clinical and theoretical perspectives. Additionally, it is well-defined, which increases its precision compared to other methods. However, this method does not account for loc al variations within the achieved plane, as it evaluates the plane as a whole. Consequently, this leads to less accurate outcomes. The dividing into several planes method is used in one study, where the achieved plane is divided into multiple straight planes, see Figure 14 (31). This method requires more steps from a theoretical perspective, making it more complex to implement. The remaining eight studies belong to the multiple point selection method. These studies select points (ranging from three to infinite points) on the achieved and planned plane, with these points assigned either randomly (15, 38, 39, 41, 48) or manually (21, 24, 32). One of these studies creates a point cloud model from the resection specimen and the planned plane, allowing for the determination of multiple distances, see Figure 15 (38). This method along with the dividing into several planes method allows for deviation measurements per point or section, which makes them capable of describing local variations within the achieved plane. Specifically the point cloud model calculates multiple distances, which is particularly beneficial to describe these local variations. However, in the multiple point selection method, the points are often not specified, resulting in a lower precision than the dividing into several planes method. For an overview of the comparisons between the methods, see Table 6. For an overview of the methods used in each study, see Table 7. For a complete overview, refer to Table 12 in Appendix B.

Figure 14: Assessment of curved planes. (A) Example of curved resection plane. (B) curved plane was divided into several adjacent straight planes. Figure copied from the study by Müller et al. (2020) (31).

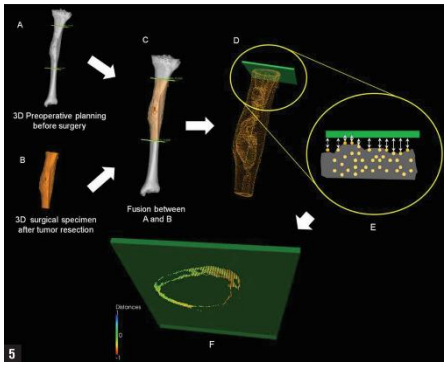


Figure 15: (A) Three-dimensional (3D) preoperative planning of the tibia. (B) 3D virtual model of the tumor surgical specimen (orange). (C) Superimposition of A and B. (D) Conversion of the virtual specimen into a point cloud. (E) Accuracy assessment by measuring point-to-point distances. Figure copied from the study by Ritacco et al. (2013) (38).

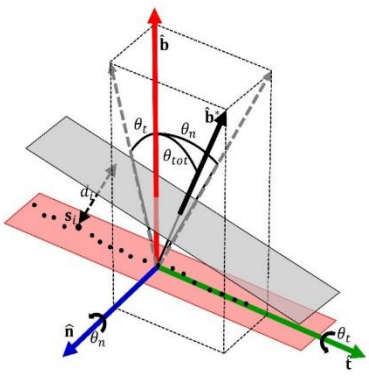


Figure 16: Selecting multiple points (s_i) in the achieved plane. Figure copied from the study by De Stavola et al. (2022) (21).

Table 6: Subjective scoring of papers based on grounders for describing the deviation within the plane. (-) indicates poor performance, (+/-) indicates average performance, (+), indicates good performance, and (++) indicates excellent performance (e.g. '+' for complexity means the method is not complex, and '+' for implementation means the method is easy to implement).

	Method 1	**Method 2**	**Method 3**	**Method 4**
	Intraoperative	Flatness parameter	Dividing into several planes	Selecting multiple points
Complexity	-	+	+	++
Implementation	+	+	+/-	+
Accuracy	+	+/-	+	+
Precision	+	++	+	+/-
Local variations	+/-	-	++	+

Table 7: Overview of the methods to determine the deviations within the achieved plane per study

Deviations within achieved plane	**First author and year**
Intraoperative	Wong, 2013
Flatness parameter	Cartiaux, 2014; Khan, 2013
Dividing achieved plane into several straight planes	Müller, 2020
Selecting multiple points in achieved plane	Bai, 2014; De Stavola, 2022; Gerbers, 2013; Otsuki, 2013; Siegel, 2020; Zhao, 2024; Sun, 2021; Ritacco, 2013

3.4 Outcome measures

Comparing two planes with each other requires both a description and quantification of their deviation. Several methods have been proposed to determine the linear and angular deviations between two planes.

3.4.1 Linear deviation

Of the 36 included studies, 34 describe linear deviations between the achieved and planned planes. Outcome measures include distance (error) and cutting accuracy.

3.4.1.1 Distance (error)

The distance (error) is categorized into five subgroups; mean distance between points, and between center of gravity (COG) or centroids, median distance between COG, and the cumulative distance between achieved and planned planes.

Among the 27 studies (79%) that use distance (error), 18 (67%) rely on points (either randomly selected or landmarks) and belong to the mean distance between points subgroup (14, 21, 23, 24, 29, 31-41, 45, 48). When multiple points are used, this subgroup accounts for local variations, making it better suited to describe these variations within the achieved plane. However, if the selected points are not well-defined, this subgroup exhibits lower precision. Four studies (15%) use the mean distance between the COG or centroids of the planes (16, 22, 27, 42), and two studies (7%) use alternative strategies (17, 30). One study belongs to the median distance between COG subgroup, which is based on the COG of the achieved and planned plane (17). This subgroup is more robust to outliers compared to the mean distance between COG subgroup. Another study belongs to the cumulative distance

subgroup, which describes the cumulative distance between the achieved and planned plane (30). This subgroup is also more robust to outliers. However it is only useful when working with a large dataset, as it attenuates the effect of individual deviations. Additionally, the point selection in this subgroup is not as well-defined as in the mean and median distance between COG and perpendicular distance subgroups, resulting in lower precision. The remaining three studies (11%) do not specify their method (15, 26, 47). Examples of the mean distance between points, and between COG or centroids subgroups are shown in Figure 15, 16 and 17. For an overview of the comparisons between the subgroups, see Table 8.

From the studies using points, four use XYZ coordinates in a digital coordinate system (33-36). One study measures distance deviation for three specific points along the planned plane compared to the achieved plane (32). The remaining studies describe distance orientation using the distance between selected points (14, 21, 23, 24, 29, 31, 37-41, 48), with one using a validated index measurement to quantify deviation (21). This index measurement consists of three displacement errors; root mean square displacement error (δ_{RMS}), which represents the overall deviation from the planned plane; signed average displacement (δ_{mean}), which discriminates between a R0, and R1 or R2 resection; and residual standard deviations (δ_b), which serves as a quality parameter for plane fitting of the achieved plane. These displacement errors were computed based on the distances d_i of points s_i on the achieved plane from the planned plane, see Figure 18.

Among the studies that do not specify their method, two describe distance orientation between planes (15, 26), while the third considers it inapplicable, as it evaluates the deviation for a single point on four sides of the tumor (47). For an overview of the methods used in each study, see Table 9. For a complete overview, refer to Table 12 in Appendix B.

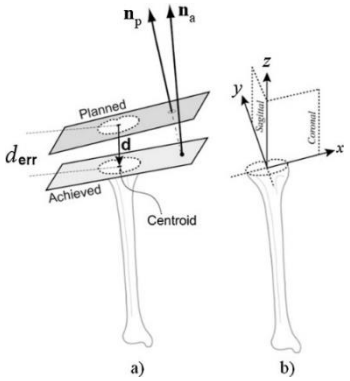


Figure 17: (A) Achieved and planned planes, defined by their normal vectors (n_p and n_a), are used to find two bone cross sections (dotted polygons). The distance error (d_{err}) is the distance between the centroids of these cross sections. (B) Vectors are projected into sagittal and coronal planes of an anatomical coordinate system, with angular differences providing the sagittal plane angulation error (θ_{err}) and coronal plane angulation error (Υ_{err}). Figure copied from the study by Dobbe et al. (2014)

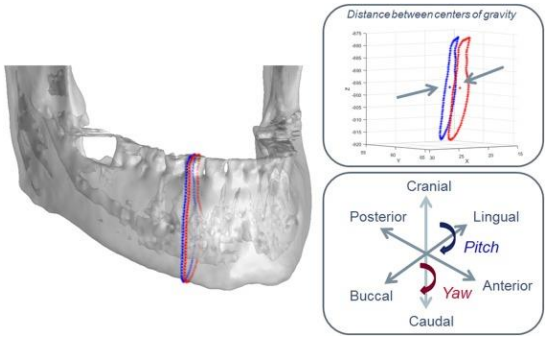


Figure 18: Achieved and planned plane delineation through the mandible, with outcomes including the distance between the planes' centers of gravity, yaw and pitch angles. Figure copied from the study by Brouwer de Koning et al. (2021) (16).

3.4.1.2 Cutting accuracy

Cutting accuracy is categorized into four subgroups; location accuracy, surgical margin (errors) relative to tumor tissue, cross-section comparison, and flatness parameter.

Among the seven studies (21%) using cutting accuracy, four studies (57%) belong to the location accuracy subgroup, using the maximum distance (18, 20, 25, 28), with three studies using ISO location accuracy standards (18, 25, 28). The fourth study evaluated location accuracy by measuring surgical margins on resection specimens, which were sectioned into two slices (20). The distance from the achieved plane to the bone cement, simulating a tumor, was measured at four locations per slice and considered the achieved surgical margin (Fig 4B). The resection error was the difference between the achieved and planned surgical margins, with the largest error used for each plane. Three of the four studies using location accuracy, evaluate cutting accuracy relative to tumor tissue (18, 20, 25), and one measures it between two planes (28). The location accuracy subgroup includes the maximum distance, which offers more insight into the amount of excessive tissue removed and provides a consistent, internationally recognized framework for evaluating cutting accuracy. This makes it useful for comparing results across studies or clinical settings and allows for a more comprehensive analysis of the achieved plane's location.

Four studies (57%) belong to the surgical margin (errors) subgroup, and use the surgical margin relative to tumor tissue (2, 18, 19, 25). Surgical margin is defined as the minimum distance between the achieved plane and the boundary of the tumor (Fig 1A.) The minimal distance is important for ensuring that all tumor tissue is removed, while the mean distance provides a more general evaluation. The surgical margin subgroup is clinically relevant as it directly relates to tumor resection margins.

Two studies (29%) belong to the flatness parameter subgroup, and thus use ISO flatness parameter standards (18, 28). One study evaluates flatness relative to tumor tissue (18), while the other assesses flatness between two planes (28). This subgroup, along with the location accuracy subgroup, has a less straightforward outcome measure, making it more complex to understand. However, it is the only outcome measure that provides information about local variations in the achieved plane.

The remaining study (14%) belongs to the cross-section comparison subgroup (Fig. 2A) (46). This study evaluates the achieved plane against the prosthesis template intraoperatively, but lacks distance orientation. This subgroup allows for a visual and direct comparison, but does not provide quantitative information. Additionally, the methodology is not well-defined, resulting in lower precision. This subgroup is also the least accurate. For an overview of the comparisons between the subgroups, see Table 8. For an overview of the methods used in each study, see Table 9.

Table 8: Subjective scoring of papers based on grounders for describing the outcome measures to determine the linear deviations between the achieved and planned plane. (-) indicates poor performance, (+/-) indicates average performance, (+), indicates good performance, and (++) indicates excellent performance (e.g. '+' for complexity means the method is not complex, and '+' for implementation means the method is easy to implement).

	Distance (error)				**Cutting accuracy**			
	S1	S2	S3	S4	S1	S2	S3	S4
	Mean distance between points	Mean distance between COG	Median distance between COG	Cumulative distance between planes	Location accuracy	Surgical margin (errors)	Cross-section comparison	Flatness parameter
Complexity	++	++	++	+	+	++	++	+
Implementation	++	++	++	+	+	++	++	+
Accuracy	+	+	+	+	+	+	-	+/-
Precision	+/-	++	++	+	++	++	-	++
Local variations	++	-	-	+	-	-	-	+

Table 9: Overview of the linear deviation outcome measures per study.

Outcome measure	**First author and year**
Mean distance (error)	
- **Landmarks or points**	Liu, 2024; Qu, 2015; Pietruski 2020; Pietruski 2019; Pietruski, 2019; Pietruski 2023; Zhao, 2024; Ackermann 2021; Sternheim, 2015; Sun, 2021; Otsuki, 2013; De Stavola, 2022; Winnand 2022; Siegel, 2020; Fujiwara, 2018; Müller, 2020; Ritacco, 2013; Gerbers, 2013
- **COG**	Brouwer de Koning, 2021; Hoving, 2018; Ter Braak, 2020; Dobbe, 2014 Modabber, 2022; Caiti, 2021
- **Alternative strategies**	Hasan, 2020; Wong, 2015; Bai, 2014
- **Method unknown**	
Cutting accuracy	
- **Location accuracy**	Gouin, 2014; Cartiaux, 2014; Khan, 2013; Cho, 2018
- **Surgical margin (error)**	Gouin, 2014; Cartiaux, 2014; Evrard, 2022; Chan 2022
- **Cross-section**	Wong, 2013
- **Flatness (ISO)**	Cartiaux, 2014; Khan, 2013

3.4.2 Angular deviations

Of the 36 included studies, 24 describe angular deviations between the achieved and planned planes(14, 16, 17, 19, 21-23, 26-30, 33-37, 40-45, 48). Outcome measures include Roll-Pitch-Yaw angles relative to the entry line of the achieved plane, and mean angle (error). The latter is determined using various methods.

Six studies (25%) use Roll-Pitch-Yaw angles as outcome measure (16, 19, 26, 28, 40, 42), see Figure 17. The 18 remaining studies (75%) use mean angle (error) as angular outcome measure, which are categorized into four subgroups: normal vector angle, centerline angle, perpendicular angle, and cumulative angle. Roll, pitch and yaw angles are more clearly specified, resulting in higher precision. Additionally, they are more suitable for detailed 3D analyses, which make them particularly useful for complex tumor resections. However, this outcome measure is more complex to understand for non-technicians.

Among the 18 studies describing mean angle (error) (75%), ten (56%) belong to the normal vector angle subgroup (17, 21, 22, 33-37, 41, 45). They calculate the angle using the normal vectors of the two planes (Fig. 17). One of these studies uses a validated index measure (21). This index measure consists of three angular errors; overall angular error (θ_{tot}), which represents the angle between the achieved and planned plane; signed around-tangent angular error (θ_t), which measures the angulation of the piezoelectric tool relative to the planned plane, distinguishing between R0, and R1 and R2 resections; and around-normal angular error (θ_n), which describes the obliquity of the incisions in the mesial-distal or cranial-apical directions, indicating differences in the resected specimen from one cut end to the other. The normal vector angle subgroup is the least complex to understand compared to the other subgroups. Furthermore, it is easier to implement from a theoretical perspective.

Three studies (17%) use alternative strategies (14, 27, 30). One study belongs to the centerline angle subgroup, which automatically generates cylinders around the bone segments with a maximum diameter and calculates the angle between the centerlines of these cylinders (27, 51) (Fig. 19). This subgroup is practical for cylindrical or anatomical straight structures. However, this method is not suitable for non-cylindrical structures. One study belongs to the perpendicular angle subgroup, which measures specific anatomical angles and compares the achieved to the planned angle (Fig. 20) (14). This subgroup provides a direct measurement in the perpendicular direction, which is particularly relevant for surgical precision. However, it is more complex to understand compared to the other subgroups. Another study belongs to the cumulative angle subgroup, and thus describes the cumulative angle (30), but does not report the plane orientation.

The remaining five studies (28%) do not describe their method and plane orientation (23, 29, 43, 44, 48). For an overview of the comparisons between the angular outcome measures and subgroups, see Table 10. For an overview of the angular outcome measures used in each study, see Table 11. For a detailed overview, see Table 12 in Appendix B.

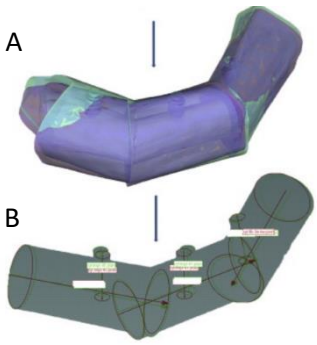


Figure 19: (A) Alignment of pre-(purple) and postoperative (blue) 3D model. (B) Automated cylinders are situated on the fibula segments to determine the center point distances and axis deviations. Figure copied from the study by Schepers et al. (2015) (54).

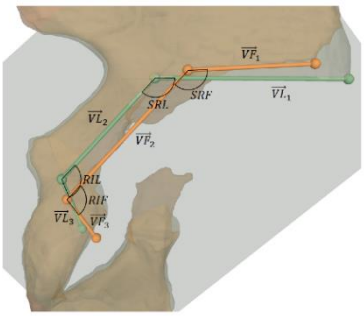


Figure 20: Projected osteotomy starting points, connecting vectors, and angles between supra- and retro-acetabular osteotomies (SRL, SRF), and between retro-acetabular and ischial osteotomies (RIL,RIF). SRL/RIL meaning planned, SRF/RIF meaning performed. Figure copied from the study by Ackermann et al. (2021) (14).

Table 10: Comparison of the outcome measures to determine the angular deviations between the achieved and planned plane. (-) indicates poor performance, (+/-) indicates average performance, (+), indicates good performance, and (++) indicates excellent performance (e.g. '+' for complexity means the method is not complex, and '+' for implementation means the method is easy to implement).

	Roll, pitch and yaw angles	**Mean angle (error)**			
	S1	S1 Angle between normal vectors	S2 Angle between centerlines	S3 Angle perpendicular to planned plane	S4 Cumulative angle
Complexity	+/-	++	+/-	-	+
Implementation	++	++	+	+	+
Accuracy	++	++	++	++	+
Precision	++	+	+	+	+
Local variations	-	-	-	-	-

Table 11: Overview of the angular deviation outcome measures per study

Outcome measure	**First author and year**
Roll-Pitch-Yaw angles	
- **Roll and pitch**	Hasan, 2020; Chan, 2022; Khan, 2013; Sternheim, 2015
- **Pitch and yaw**	Brouwer de Koning, 2021
- **Roll and yaw**	Ter Braak, 2020
Mean angle (error)	
- **Normal vectors**	De Stavola, 2022; Qu, 2015; Pietruski, 2019; Pietruski, 2019; Pietruski, 2020; Pietruski, 2023; Winnand, 2022; Caiti, 2021; Dobbe, 2014; Sun, 2021
- **Alternative strategies**	Hoving, 2018; Ackermann, 2021; Modabber, 2022
- **Unknown**	Fujiwara, 2018; Liu, 2024; Vanesa, 2021; Weijs, 2016; Zhao, 2024

3.4.3 Oncology

Of the included studies, 19 have an oncological focus (2, 15, 18-20, 23-29, 31, 38-40, 46-48). When determining the outcome measures related to tumor tissue, other parameters become important, with the surgical margin being the most critical clinical outcome measure. Therefore, the achieved plane should be cut further away from the tumor rather than closer, see Figure 21. Additionally, while volume is not an important outcome measure in correction osteotomies for assessing the difference between achieved and planned planes, it may be a useful outcome measure for resection specimens, as it provides an initial indication of the outcome. Only one study with an oncological focus has already used volume as a postoperative outcome measure (48).

Seven studies (37%) use cutting accuracy as linear outcome measure (2, 18-20, 25, 28, 46), while twelve studies (63%) use distance (error) (15, 23-27, 29, 31, 38-40, 47, 48). Of the latter group, eight (67%) rely on points (23, 24, 29, 31, 38-40, 48). The study of Gerbers et al. (2013) determined the mean gap distance between points on the host and donor bone (24), while the studies of Fujiwara et al. (2018) and Sternheim et al. (2015) determined the mean distance between entry and exit points (23, 40). The study of Liu et al. (2024) and determined the mean distance using landmark points (29), and the remaining five studies determined the mean distance using randomly selected points (29, 31, 38, 39, 48). One out of 12 studies (8%) uses the mean distance between the COG of the planes (27), and they remaining three studies (25%) do not specify the method (15, 26, 47)

Nine out of 19 studies determine the margin using histological laboratory tests (2, 15, 20, 23-25, 31, 46, 47) ,while seven studies use software for this purpose (18, 19, 27, 28, 38-40). The remaining four studies do not conduct margin assessments (15, 26, 29, 48).

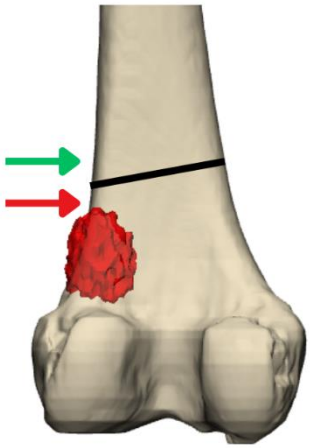


Figure 21: Bone tumor (red). Black line indicates planned plane. Green arrow indicates proximal area where achieved plane should be made. Red arrow indicates distal area where achieved plane should not be made (e.g. closer to tumor).

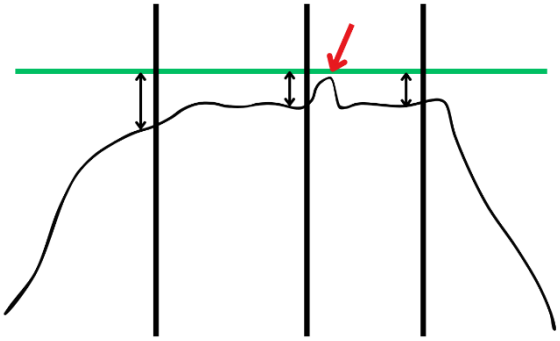


Figure 22: Missed area of tumor involvement. Green line indicates achieved plane. Black lines indicate selected section of the margin determined by pathologists. Red arrow indicates missed area of tumor involvement.

Histological laboratory tests are considered the gold standard for surgical margin determination due to their high precision and accuracy (52). They also provide valuable insights into tumor biology, including cellular atypia and invasion patterns (53). However, these tests are time-intensive and require specialized equipment and trained pathologists. Additionally, only selected coupes of the margin are examined, which may result in missed areas of tumor involvement (54), see Figure 22. In contrast, software-based margin determination offers a faster alternative and enables more quantitative analysis through the use of algorithms, thereby reducing subjectivity. Additionally, it allows for the analysis of an unlimited number of sections, potentially addressing the question of whether a resection plane measuring 10 cm with a single 1 mm margin peak poses the same risk as a 10 cm resection plane with a consistent 1 mm margin throughout (Fig. 23). However, its accuracy is highly dependent on the quality and resolution of the imaging modality used and lacks the ability to provide detailed tissue characterization.

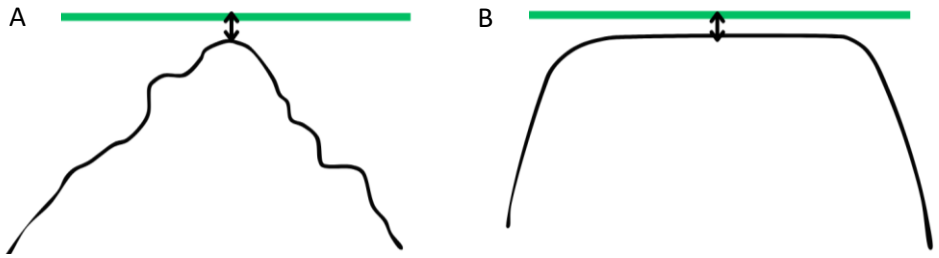


Figure 23: Surgical margin between achieved plane of 10 cm (green) and resection specimen. (A) single peak with 1 mm margin, (B) consistent 1 mm margin. The cases potentially have different risks for local recurrence.

Eight studies with an oncological focus describe angular deviations. Fifty percent of these studies use Roll-Pitch-Yaw angles as the outcome measure (19, 26, 28, 40). The remaining 50% describe the mean angle (error) (23, 27, 29, 48). Within the mean angle (error) outcome measure, one study

automatically generates cylinders around the bone segments with a maximum diameter and calculates the angle between the centerlines of these cylinders (27, 51) (Fig. 19). The remaining three studies do not specify their methodology and plane orientation (23, 29, 48).

4.0 Discussion

The aim of this scoping review was to summarize the methods used to describe the achieved plane, and the associated linear and angular outcome measures, in order to establish a more standardized approach for evaluating the procedural accuracy of resected bone (tumor) compared to the preoperative plan. We identified 36 primary studies addressing the linear and/or angular deviation between achieved and planned planes, published between 2013 and 2024. Our findings indicate that there is a lack of standardized methods and outcome measurements in this field.

Most of the studies do not specify how the achieved plane is determined. While several 2D and 3D methods are described, 3D measurements provide a more accurate description of the achieved plane. 3D methods with low complexity, easy implementation from both clinical and theoretical perspectives, high accuracy and precision, and the ability to describe local variations are promising.

The point cloud method of Ritacco et al. (2013) (38), the mesh portions method of Modabber et al. (2022) (30) and the normal vector method of Brouwer de Koning et al. (2021) (16) have the most potential to describe the achieved plane according to these parameters (see Table 4). Future research will determine which method is best for standardization. Additionally, deviations within the achieved plane are often not specified, despite their importance. The dividing into several planes method of Müller et al. (2020) (31), and the point cloud method of Ritacco et al. (2013) (38) have the most potential to describe local variations within the achieved plane according to the same parameters (see Table 6). Future research will determine which method is best for standardization.

Before the difference between the achieved and planned planes can be quantified, postoperative CT and/or MRI scans are required. Nine studies scanned resection specimens, while the others scanned patients. The primary advantage of the former method is that it eliminates additional radiation exposure to the patient. To determine the difference between the planes, the outcome measure should be corrected for the thickness of the saw blade to account for the loss of bone material (kerf) during resection (Fig. 1B). Most studies do not specify whether they accounted for kerf, although cutting loss can influence the dimensions of the resected specimen, leading to an underestimation of the postoperative tumor size. Additionally, cutting loss can affect the surgical margin, both of which have implications for follow-up treatment.

When examining the outcome measures in studies describing osteotomy planes, most of the studies use the distance and angle (error). 63% of them use landmarks or points to determine the linear deviation and 56% use normal vectors of the achieved and planned planes to determine the angular deviation.

When examining the outcome measures in studies with an oncological focus, most studies use distance (error) (63%) as well, with 67% focusing on mean distance between the achieved and planned planes. The remaining studies (37%) use cutting accuracy as outcome measure. Notably, all studies using cutting accuracy focus specifically on bone tumor resections. Eight studies determine the margin using histological laboratory tests, while seven studies use software for this purpose. The studies are

equally used on angular outcome measures, but those using mean angle (error) do not specify the orientation of the angles.

Distance (error) outcome measures, including the mean distance between points, and the mean and median distance between COG, as well as cutting accuracy outcome measures, including surgical margin (errors) and , have the most potential based on the established parameters. Mean angle (error) outcome measures, including angle between normal vectors of the achieved and planned plane, and roll, pitch and yaw angles have the most potential based on these same parameters. Future research will determine which linear and angular method is best for standardization.

Our scoping review has some limitations. Firstly, identifying all relevant articles through a search is challenging. Secondly, the included studies use a variety of methodologies and outcome measures. Although these studies have been grouped into subgroups based on their closest relevance, this approach may lead to an inaccurate overall summary. Thirdly, several assumptions were made to address missing details from methods, such as the determination of the achieved plane, deviations within the plane, and the orientation used. These issues were discussed with a biomechanical engineer (B.K) and conclusions were drawn accordingly. However, these assumptions reduce our confidence in the result. Lastly, no quality assessment was performed on the included articles, which could potentially influence the overall findings.

5.0 Conclusion

In conclusion, there is a lack of standardized methods and outcome measures for determining the linear and angular deviations between the achieved and planned planes in bone (tumor) resections. Moreover, the majority of studies do not clearly describe their methodologies. Future research is needed to identify the most suitable methods and outcome measures for standardization.

References

1. Portnoy Y, Koren J, Khoury A, Factor S, Dadia S, Ran Y, et al. Three-dimensional technologies in presurgical planning of bone surgeries: current evidence and future perspectives. International journal of surgery. 2023.
2. Evrard R, Schubert T, Paul L, Docquier PL. Quality of resection margin with patient specific instrument for bone tumor resection. J Bone Oncol. 2022;34:100434.
3. Hammoudeh JA, Howell LK, Boutros S, Scott MA, Urata MM. Current Status of Surgical Planning for Orthognathic Surgery: Traditional Methods versus 3D Surgical Planning. Plastic and reconstructive surgery Gloabl open. 2015.
4. Wade RH, Kevu J, J. D. Pre-operative planning in orthopaedics: a study of surgeons' opinions. Injury. 1998.
5. Mezger U, Jendrewski C, Bartels M. Navigation in surgery. Langenbeck's archives of surgery. 2013.
6. Bell R. Computer planning and intraoperative navigation in cranio-maxillofacial surgery. Oral and maxillofacial surgery clinics of North America. 2010.
7. Sallent A, Vicente M, Reverté MM, Lopez A, Rodríguez-Baeza A, Pérez-Domínguez M, et al. How 3D patient-specific instruments improve accuracy of pelvic bone tumour resection in a cadaveric study. Bone & joint research. 2017.
8. Benignus C, Buschner P, Meier M, Wilken F, Rieger J, Beckmann J. Patient Specific Instruments and Patient Individual Implants-A Narrative Review. Journal of personalized medicine. 2023.
9. Han S-B, Kim H-J, Kim T-K, In Y, Oh K-J, Koh I-J, et al. Computer navigation is effective in reducing blood loss but has no effect on transfusion requirement following primary total knee arthroplasty: a meta-analysis. Knee Surgery, Sports Traumatology, Arthroscopy. 2016;24(11):3474-81.
10. Wong KC, Sze LKY, Kumta SM. Complex joint-preserving bone tumor resection and reconstruction using computer navigation and 3D-printed patient-specific guides: A technical note of three cases. . Journal of orthopaedic translation. 2021.
11. Brown CL, Marcaccio S, Mullen JP, Lin RT, McMahon S, Karimi A, et al. Patient Specific Instrumentation in Shoulder Arthroplasty. Operative Techniques in Orthopaedics. 2024;34(4):101152.
12. Stiehl JB. Computer-Assisted Surgery: Pros and Cons. In: Scuderi GR, Tria AJ, editors. Minimally Invasive Surgery in Orthopedics. Cham: Springer International Publishing; 2016. p. 1-9.
13. Ma L, Zhou Y, Zhu Y, Lin Z, Wang Y, Zhang Y, et al. 3D-printed guiding templates for improved osteosarcoma resection. Scientific Reports. 2016;6(1):23335.
14. Ackermann J, Liebmann F, Hoch A, Snedeker JG, Farshad M, Rahm S, et al. Augmented Reality Based Surgical Navigation of Complex Pelvic Osteotomies-A Feasibility Study on Cadavers. Appl Sci-Basel. 2021;11(3):19.
15. Bai G, He D, Yang C, Lu C, Huang D, Chen M, et al. Effect of digital template in the assistant of a giant condylar osteochondroma resection. J Craniofac Surg. 2014;25(3):e301-4.
16. Brouwer de Koning SG, Ter Braak TP, Geldof F, van Veen RLP, van Alphen MJA, Karssemakers LHE, et al. Evaluating the accuracy of resection planes in mandibular surgery using a preoperative, intraoperative, and postoperative approach. Int J Oral Maxillofac Surg. 2021;50(3):287-93.
17. Caiti G, Dobbe JGG, Strackee SD, van Doesburg MHM, Strijkers GJ, Streekstra GJ. A 3D printed cast for minimally invasive transfer of distal radius osteotomy: a cadaver study. Int J Comput Assist Radiol Surg. 2021;16(3):505-13.
18. Cartiaux O, Paul L, Francq BG, Banse X, Docquier PL. Improved accuracy with 3D planning and patient-specific instruments during simulated pelvic bone tumor surgery. Ann Biomed Eng. 2014;42(1):205-13.
19. Chan HHL, Sahovaler A, Daly MJ, Ferrari M, Franz L, Gualtieri T, et al. Projected cutting guides using an augmented reality system to improve surgical margins in maxillectomies: A preclinical study. Oral Oncology. 2022;127:7.
20. Cho HS, Park MS, Gupta S, Han I, Kim HS, Choi H, et al. Can Augmented Reality Be Helpful in Pelvic Bone Cancer Surgery? An In Vitro Study. Clin Orthop Relat Res. 2018;476(9):1719-25.

21. De Stavola L, Cristoforetti A, Fincato A, Nollo G, Ghensi P, Cantarutti A, et al. Accuracy and Technical Predictability of Computer Guided Bone Harvesting from the Mandible: A Cone-Beam CT Analysis in 22 Consecutive Patients. J Funct Biomater. 2022;13(4).
22. Dobbe JG, Kievit AJ, Schafroth MU, Blankevoort L, Streekstra GJ. Evaluation of a CT-based technique to measure the transfer accuracy of a virtually planned osteotomy. Med Eng Phys. 2014;36(8):1081-7.
23. Fujiwara T, Kunisada T, Takeda K, Hasei J, Nakata E, Nakahara R, et al. Intraoperative O-arm-navigated resection in musculoskeletal tumors. J Orthop Sci. 2018;23(6):1045-50.
24. Gerbers JG, Ooijen PM, Jutte PC. Computer-assisted surgery for allograft shaping in hemicortical resection: a technical note involving 4 cases. Acta Orthopaedica. 2013.
25. Gouin F, Paul L, Odri GA, Cartiaux O. Computer-Assisted Planning and Patient-Specific Instruments for Bone Tumor Resection within the Pelvis: A Series of 11 Patients. Sarcoma. 2014;2014:842709.
26. Hasan W, Daly MJ, Chan HHL, Qiu J, Irish JC. Intraoperative cone-beam CT-guided osteotomy navigation in mandible and maxilla surgery. Laryngoscope. 2020;130(5):1166-72.
27. Hoving AM, Kraeima J, Schepers RH, Dijkstra H, Potze JH, Dorgelo B, et al. Optimisation of three-dimensional lower jaw resection margin planning using a novel Black Bone magnetic resonance imaging protocol. PLoS One. 2018;13(4):e0196059.
28. Khan F, Pearle A, Lightcap C, Boland PJ, Healey JH. Haptic robot-assisted surgery improves accuracy of wide resection of bone tumors: a pilot study. Clin Orthop Relat Res. 2013;471(3):851-9.
29. Liu Z, Zhong Y, Lyu X, Zhang J, Huang M, Liu S, et al. Accuracy of the modified tooth-supported 3D printing surgical guides based on CT, CBCT, and intraoral scanning in maxillofacial region: A comparison study. J Stomatol Oral Maxillofac Surg. 2024;125(5s2):101853.
30. Modabber A, Ayoub N, Redick T, Gesenhues J, Kniha K, Möhlhenrich SC, et al. Comparison of augmented reality and cutting guide technology in assisted harvesting of iliac crest grafts - A cadaver study. Ann Anat. 2022;239:151834.
31. Müller DA, Stutz Y, Vlachopoulos L, Farshad M, Fürnstahl P. The Accuracy of Three-Dimensional Planned Bone Tumor Resection Using Patient-Specific Instrument. Cancer Manag Res. 2020;12:6533-40.
32. Otsuki B, Takemoto M, Kawanabe K, Awa Y, Akiyama H, Fujibayashi S, et al. Developing a novel custom cutting guide for curved peri-acetabular osteotomy. Int Orthop. 2013;37(6):1033-8.
33. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Jaworowski J, et al. Supporting fibula free flap harvest with augmented reality: A proof-of-concept study. Laryngoscope. 2020;130(5):1173-9.
34. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Mazurek M, et al. Supporting mandibular resection with intraoperative navigation utilizing augmented reality technology - A proof of concept study. J Craniomaxillofac Surg. 2019;47(6):854-9.
35. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Mazurek M, et al. Navigation-guided fibula free flap for mandibular reconstruction: A proof of concept study. J Plast Reconstr Aesthet Surg. 2019;72(4):572-80.
36. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Świecka M, et al. Replacing cutting guides with an augmented reality-based navigation system: A feasibility study in the maxillofacial region. Int J Med Robot. 2023;19(3):e2499.
37. Qu M, Hou Y, Xu Y, Shen C, Zhu M, Xie L, et al. Precise positioning of an intraoral distractor using augmented reality in patients with hemifacial microsomia. J Craniomaxillofac Surg. 2015;43(1):106-12.
38. Ritacco LE, Milano FE, Farfalli GL, Ayerza MA, Muscolo DL, Aponte-Tinao LA. Accuracy of 3-D planning and navigation in bone tumor resection. Orthopedics. 2013;36(7):e942-50.
39. Siegel MA, Balach T, Sweeney KR, Nystrom LM, Colman MW. Sacroiliac joint cut accuracy: Comparing new technologies in an idealized sawbones model. J Surg Oncol. 2020;122(6):1218-25.

40. Sternheim A, Daly M, Qiu J, Weersink R, Chan H, Jaffray D, et al. Navigated pelvic osteotomy and tumor resection: a study assessing the accuracy and reproducibility of resection planes in Sawbones and cadavers. J Bone Joint Surg Am. 2015;97(1):40-6.
41. Sun M, Lin L, Chen X, Xu C, Zin MA, Han W, et al. Robot-assisted mandibular angle osteotomy using electromagnetic navigation. Ann Transl Med. 2021;9(7):567.
42. Ter Braak TP, Brouwer de Koning SG, van Alphen MJA, van der Heijden F, Schreuder WH, van Veen RLP, et al. A surgical navigated cutting guide for mandibular osteotomies: accuracy and reproducibility of an image-guided mandibular osteotomy. Int J Comput Assist Radiol Surg. 2020;15(10):1719-25.
43. Vanesa V, Irene MP, Marta AS, Francisco José PF, Miguel BS, Mireia RM, et al. Accuracy of virtually planned mandibular distraction in a pediatric case series. J Craniomaxillofac Surg. 2021;49(2):154-65.
44. Weijs WL, Coppen C, Schreurs R, Vreeken RD, Verhulst AC, Merkx MA, et al. Accuracy of virtually 3D planned resection templates in mandibular reconstruction. J Craniomaxillofac Surg. 2016;44(11):1828-32.
45. Winnand P, Ayoub N, Redick T, Gesenhues J, Heitzer M, Peters F, et al. Navigation of iliac crest graft harvest using markerless augmented reality and cutting guide technology: A pilot study. Int J Med Robot. 2022;18(1):e2318.
46. Wong KC, Kumta SM. Joint-preserving tumor resection and reconstruction using image-guided computer navigation. Clin Orthop Relat Res. 2013;471(3):762-73.
47. Wong KC, Kumta SM, Geel NV, Demol J. One-step reconstruction with a 3D-printed, biomechanically evaluated custom implant after complex pelvic tumor resection. Comput Aided Surg. 2015;20(1):14-23.
48. Zhao Z, Zhang Y, Lin L, Huang W, Xiao C, Liu J, et al. Intelligent electromagnetic navigation system for robot-assisted intraoral osteotomy in mandibular tumor resection: a model experiment. Front Immunol. 2024;15:1436276.
49. Tricco A, Lillie E, Zarin W, O'Brien K, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Annals of Internal Medicine. 2018;169(7):467-73.
50. 5725-1 I. Accuracy (trueness and precision) of measurement methodsand results–Part 1: General principles and definitions. ISO. Geneva2023.
51. Schepers RH, Raghoebar GM, Vissink A, Stenekes MW, Kraeima J, Roodenburg JL, et al. Accuracy of fibula reconstruction using patient-specific CAD/CAM reconstruction plates and dental implants: A new modality for functional reconstruction of mandibular defects. Journal of Cranio-Maxillofacial Surgery. 2015;43(5):649-57.
52. Tseng LJ, Matsuyama A, MacDonald-Dickinson V. Histology: The gold standard for diagnosis? The Canadian veterinary journal. 2023.
53. Rodriguez-Canales J, Eberle FC, Jaffe ES, Emmert-Buck MR. Why is it crucial to reintegrate pathology into cancer research?. BioEssays : news and reviews in molecular, cellular and developmental biology. 2011.
54. Sarode G, Sarode SC, Shelke P, Patil S. Histopathological assessment of surgical margins of oral carcinomas and related shrinkage of tumour. Translational Research in Oral Oncology. 2017;2:2057178X17708078.

Appendix

A – Search term PubMed

(("cutting plane"[tw] OR "cutting planes"[tw] OR "cut plane"[tw] OR "cut planes"[tw] OR "cutting plane"[title/abstract:~0] OR "cutting planes"[title/abstract:~0] OR "cut plane"[title/abstract:~0] OR "cut planes"[title/abstract:~0] OR "sawing plane"[title/abstract:~0] OR "sawing planes"[title/abstract:~0] OR "sawe plane"[title/abstract:~0] OR "sawe planes"[title/abstract:~0] OR "sawed plane"[title/abstract:~0] OR "sawed planes"[title/abstract:~0] OR "cutting plane"[title/abstract:~2] OR "cutting planes"[title/abstract:~2] OR "cut plane"[title/abstract:~2] OR "cut planes"[title/abstract:~2] OR "sawing plane"[title/abstract:~2] OR "sawing planes"[title/abstract:~2] OR "sawe plane"[title/abstract:~2] OR "sawe planes"[title/abstract:~2] OR "sawed plane"[title/abstract:~2] OR "sawed planes"[title/abstract:~2] OR "cutting guides"[tw] OR "cutting guide"[tw] OR "cut guides"[tw] OR "cut guide"[tw] OR "cutting guides"[title/abstract:~2] OR "cutting guide"[title/abstract:~2] OR "cut guides"[title/abstract:~2] OR "cut guide"[title/abstract:~2] OR "resection planes"[tw] OR "resection plane"[tw] OR "planes resection"[title/abstract:~6] OR "planes resect"[title/abstract:~6] OR "planes resected"[title/abstract:~6] OR "planes resecting"[title/abstract:~6] OR "plane resection"[title/abstract:~6] OR "plane resect"[title/abstract:~6] OR "plane resected"[title/abstract:~6] OR "plane resecting"[title/abstract:~6]) AND ("Surgery, Computer-Assisted"[Mesh:noexp] OR "Computer Aided Surgery"[tw] OR "Computer Aided Surg*"[tw] OR "Computer Assisted Surgery"[tw] OR "Computer Assisted Surg*"[tw] OR "Image Guided Surgery"[tw] OR "Image Guided Surg*"[tw] OR "Surgical Navigation"[tw] OR "Computer-Aided Design"[mesh] OR "Computer Aided Surgery"[title/abstract:~6] OR "Computer Assisted Surgery"[title/abstract:~6] OR "Image Guided Surgery"[title/abstract:~6] OR "Surgical Navigation"[title/abstract:~6] OR "Patient-specific instrument"[tw] OR "Patient-specific instruments"[tw] OR "PSI"[tw] OR "Patient-specific instrument"[title/abstract:~6] OR "Patient-specific instruments"[title/abstract:~6] OR "Computer Aided"[tw] OR "Computer Assisted"[tw] OR "Computer Aided"[title/abstract:~6] OR "Computer Assisted"[title/abstract:~6] OR "patient specific"[tw]) AND ("Osteotomy"[Mesh] OR "osteotomy"[tw] OR "osteotom*"[tw] OR "tumor resection"[tw] OR "tumour resection"[tw] OR "tumor resection"[title/abstract:~6] OR "tumour resection"[title/abstract:~6] OR "tumors resection"[title/abstract:~6] OR "tumours resection"[title/abstract:~6] OR "tumor resected"[title/abstract:~6] OR "tumour resected"[title/abstract:~6] OR "tumors resected"[title/abstract:~6] OR "tumours resected"[title/abstract:~6]) AND ("Tomography, X-Ray Computed"[Mesh] OR "CT"[tw] OR "Computed Tomography"[tw] OR "Computer Tomography"[tw] OR "CAT Scan"[tw] OR "CAT Scans"[tw] OR "Computerized Tomography"[tw] OR "Computerised Tomography"[tw] OR "Computer Assisted Tomography"[tw] OR "Computed Tomogra*"[tw] OR "Computer Tomogra*"[tw] OR "Computerized Tomogra*"[tw] OR "Computerised Tomogra*"[tw] OR "Computer Assisted Tomogra*"[tw] OR "Magnetic Resonance Imaging"[Mesh:noexp] OR "Magnetic Resonance Imaging"[tw] OR "MRI"[tw] OR "MR Imag*"[tw] OR "Magnetic Resonance imag*"[tw]) AND (english[la] OR dutch[la]) AND ("accuracy"[tw] OR "validation"[tw] OR "evaluation"[tw] OR "accurate"[tw] OR "inaccurate"[tw] OR "accura*"[tw] OR "inaccura*"[tw] OR "validate"[tw] OR "validated"[tw] OR "validat*"[tw] OR "evaluation"[tw] OR "evaluat*"[tw] OR "Evaluation Study"[pt])

B – Content of studies

Table 12: Overview of the content of the studies. Yellow indicates that the specific topic is not determined in the study

															Achieved plane								Angular outcome measures											
First author and year	**Type of study**				**Scanned subject**		**Preoperative scan**			**Postoperative scan**					**2D**	**3D**			**Deviations within plane**				**Distance (error)**				**Cutting accuracy**				**XYZ angles**	**Mean angle (error)**		
	Clinical	Saw bones	Cadaveric	Animal	Patient	Resection specimen	CT	MRI	High-resolution laser	CT	MRI	High-resolution laser	Not Applicable	Unknown	Intraoperative	Unknown	Selecting points and plane fit	Alternative strategies	Unknown	Flatness	Selecting points	Alternative strategies	Points or landmarks	COG	Unknown	Alternative strategies	Location accuracy	Surgical margin (errors)	Flatness (ISO)	Cross-section	Roll&Pitch&Yaw	Unknown	Normal vectors	Alternative strategies
Qu, 2015	X				X		X			X						X			X				X										X	
Otsuki, 2013		X			X		X			X						X					X		X											
Brouwer de Koning, 2021	X				X		X			X							X		X					X							X			
De Stavola, 2022	X				X		X			X						X					X		X										X	
Vanesa, 2021	X				X		X			X							X		X													X		
Weijs, 2016	X				X		X			X							X		X													X		
Fujiwara, 2018	X				X		X			X								X	X				X									X		
Hasan, 2020	X				X		X			X							X		X						X						X			
Hoving, 2018	X				X		X	X		X						X			X					X										X
Liu, 2024	X				X		X			X							X		X				X									X		
Gouin, 2014	X				X		X	X		X								X	X								X	X						
Bai, 2014	X				X		X			X						X					X				X									
Gerbers, 2013	X				X		X			X						X					X		X											
Evrard, 2022	X					X	X	X			X					X			X									X						
Müller, 2020	X					X	X	X		X						X						X	X											
Wong, 2013	X					X	X	X					X		X							X								X				
Ritacco 2013	X					X	X	X		X							X				X		X											
Wong, 2015	X					X	X			X						X			X						X									

Pietruski, 2019	X			X		X		X					X			X			X								X	
Pietruski, 2019	X			X		X		X					X			X			X								X	
Pietruski, 2020	X			X		X		X					X			X			X								X	
Pietruski, 2023	X			X		X		X					X			X			X								X	
Ter Braak, 2020	X			X		X		X						X		X				X					X			
Winnand, 2022	X				X	X		X						X		X					X						X	
Chan, 2022	X			X		X					X		X			X							X		X			
Khan, 2013	X			X			X		X					X			X					X		X	X			
Cartiaux, 2014	X			X		X		X							X		X					X	X	X				
Siegel, 2020	X			X		X		X					X					X	X									
Zhao, 2024	X				X	X		X					X					X	X							X		
Ackermann, 2021		X		X		X		X						X		X			X									X
Dobbe, 2014		X		X		X		X						X		X				X							X	
Caiti, 2021		X		X		X		X						X		X					X						X	
Modabber, 2022		X			X	X		X						X		X					X							X
Sternheim, 2015		X		X		X		X						X		X			X						X			
Cho, 2018					X	X				X		X				X						X						
Sun, 2021			X	X		X		X					X					X	X								X	

C – Content oncological studies

Table 13: Overview of the content of the studies that describe tumor lesions. Yellow indicates that the specific topic is not determined in the study

															Achieved plane								Angular outcome measures											
First author and year	**Type of study**				**Scanned subject**		**Preoperative scan**			**Postoperative scan**					**2D**	**3D**			**Deviations within plane**				**Distance (error)**				**Cutting accuracy**				**XYZ angles**	**Mean angle (error)**		
	Clinical	Saw bones	Cadaveric	Animal	Patient	Resection specimen	CT	MRI	High-resolution laser	CT	MRI	High-resolution laser	Not Applicable	Unknown	Intraoperative	Unknown	Selecting points and plane fit	Alternative strategies	Unknown	Flatness	Selecting points	Alternative strategies	Points or landmarks	COG	Unknown	Alternative strategies	Location accuracy	Surgical margin (errors)	Flatness (ISO)	Cross-section	Roll&Pitch&Yaw	Unknown	Normal vectors	Alternative strategies
Fujiwara, 2018	X				X		X			X								X	X				X									X		
Hasan, 2020	X				X		X			X							X		X						X						X			
Hoving, 2018	X				X		X	X		X						X			X					X										X
Liu, 2024	X				X		X			X							X		X				X									X		
Gouin, 2014	X				X		X	X		X								X	X								X	X						
Bai, 2014	X				X		X			X						X					X				X									
Gerbers, 2013	X				X		X			X						X					X		X											
Evrard, 2022	X					X	X	X			X					X			X									X						
Müller, 2020	X					X	X	X		X						X						X	X											
Wong, 2013	X					X	X	X					X		X							X								X				
Ritacco 2013	X					X	X	X		X							X				X		X											
Wong, 2015	X					X	X			X						X			X						X									
Chan, 2022		X			X		X							X		X			X									X			X			
Khan, 2013		X			X				X			X					X			X							X			X	X			
Cartiaux, 2014		X			X		X			X								X		X							X	X		X				
Siegel, 2020		X			X		X			X						X					X		X											
Zhao, 2024		X				X	X			X						X					X		X									X		
Sternheim, 2015			X		X		X			X							X		X				X								X			
Cho, 2018						X	X						X		X				X								X							

[This page intentionally left blank]

II

TM MSc Thesis

Achieved accuracy of 3D surgical plans in complex orthopedic procedures using computer assisted surgery or patient specific instruments

Abstract

BACKGROUND: Preoperative three-dimensional (3D) planning has become the standard of care for complex orthopedic procedures. Quantitatively evaluating whether the preoperative plan is successfully performed remains a challenge due to the lack of standardized measurements, which makes comparison across studies difficult. Therefore, standardized measurements are required to close the feedback loop to improve upon our procedures and achieve safer or closer resections.

OBJECTIVE: This present study investigates the best suitable standardized approach for evaluating the procedural accuracu of the achieved plane compared to the preoperative plan, and to assess the achieved results in complex orthopedic procedures, with a main focus on bone tumor resections.

DESIGN: This single-centre retrospective cohort study included 10 patients who underwent complex orthopaedic procedures (7 oncological, 3 non-oncological), between 2021 and 2024, with a total of 28 cutting planes in patients and 17 in allografts. All patients had a 3D planned surgical approach with digital visualization and patient-specific instruments (PSIs). The achieved planes were determined using four methods (mesh, normal vector, manual and point cloud) and the procedural accuracy was assessed by multiple linear and angular outcome measures (distance between points, center of grafity (COG)-to-COG, COG-to-plane, angle between normal vectors and pitch and roll angles).

RESULTS: The mesh and point cloud methods were superior in terms of ease of use and objectivity, with the point cloud method being the most accessible due to the lack of segmentation requirements. The point cloud method performed comparably to the mesh method for linear and angular deviations, with a mean deviation within 0.8 mm (millimeters) and 0.2 degrees for roll and pitch angles in patient cases. The average flatness according to the corrected ISO-1101 standard was 2.5 mm, and a surgical margin difference of 1 mm was observed. Significant differences were observed between the COG-to-COG approach and other linear approaches for both methods, and between the corrected and non-corrected ISO-standards for the mesh method.

CONCLUSION: The point cloud method appeared to be a good alternative to the mesh method for identifying the achieved plane in bone (tumor) resections. Furthermore, the average outcomes observed in this study provide a useful baseline for assessing procedural accuracy in complex orthopaedic procedures using 3D planning and computer assisted surgery (CAS). Further research with larger sample sizes is needed to validate these findings.

KEYWORDS: bone resection, computer-assisted surgery, linear deviation, angular deviation, cutting plane, procedural accuracy

1.0 Introduction

Preoperative planning of cutting planes in orthopedic surgeries has improved procedural accuracy and increased the surgeon's efficiency, leading to reduced operation time, complication rates, and postoperative hospitalization periods (1-3). Preoperative planning is defined as the digital preparation phase prior to the actual surgery. It consists of multiple steps; with the main components including visualization, modeling, analysis, and plan generation (4). Visualization involves presenting the original two-dimensional (2D) images, including the structures of interest. Modeling consists of creating three-dimensional (3D) representations of the structures of interest. Analysis involves exploring the surgical options through simulation of the model, and plan generation includes selecting and designing the most appropriate solution for the surgical procedure based on the results of the visualization and analysis.

Conventional methods typically use a 2D image representation obtained from radiography, computed tomography (CT), magnetic resonance imaging (MRI) or ultrasound to plan the cutting plane (5). CT imaging later allowed for 3D image reconstruction through volume rendering (5), providing a more comprehensive understanding of the patient's anatomy. Currently, preoperative 3D planning has become standard of care in complex orthopedic procedures, such as multiplanar osteotomies and complex tumor resections. Accurate preoperative planning can be achieved through the use of 3D reconstructed bone models (6-8). These models are derived from CT, sometimes fused with MRI, and allow the surgeon to visualize the anatomy in 3D and virtually plan the procedure (5-8). It has already demonstrated significant value in joint preserving resections and complex reconstruction cases (6, 8). Similar practices are also in use in orthognathic or neurosurgical preoperative planning (7). Several techniques have been described to perform these plans, including navigation systems and patient-specific instrumentation (PSI) (9-13). Similar outcomes have been reported for these techniques (14). Both offer advantages over conventional methods, including reduced surgical time, exposure and blood loss, as well as improved accuracy, resulting in better surgical margins and patient outcomes (14-16).

Complex orthopedic procedures require high precision and accuracy. In the context of malignant bone tumors, a radical resection with safe tumor margins is the most critical prognostic factor that the surgeon can influence for a successful outcome, as it is strongly associated with a low local recurrence rate (17-24). However, excessive resection of healthy tissue can compromise reconstruction options and may lead to poor functional outcomes (21). For example, joint-replacing reconstructions (e.g. tumor prosthesis) are required, instead of joint salving (e.g. Capanna type intercalary reconstruction). The joint-replacing reconstructions have more limited durability (25).

Orthopedic oncology continues to push its boundaries by striving to resect larger and complex tumors with more complex preoperative plans aiming for negative margins, while preserving adjacent critical structures to maintain postoperative function (26). Therefore, precise knowledge of the tumor's involvement is essential prior to surgery, needing an accurate planning (11, 19). In this context, surgical margin is the crucial factor for maintaining a safe distance from the lesion while persevering as much healthy tissue as possible (27).

Many studies have investigated the deviations between the 3D planned and achieved planes (6, 11, 19, 23, 26-57). Quantitatively evaluating whether the preoperative plan is successfully performed

remains a challenge due to the lack of standardized measurements. These are however required to close the feedback loop to improve upon our procedures and achieve safer or closer resections.

As preliminary research, a comparative literature review was conducted (58). In this review, various methodologies and outcome measures for determining the procedural accuracy of achieved planes were evaluated. Thirty-six articles were assessed (6, 11, 19, 23, 26-57). These articles used a wide range of methods, many of which were incomplete and/or poorly described, particularly in terms of achieved plane determination, local variations within this plane, and the assessment of linear and angular outcome measures. It can be concluded that a standardized and well-described methodology for evaluating procedural accuracy by comparing deviations between planned and achieved planes is imperative.

Therefore, the purpose of this thesis is to investigate the best suitable standardized approach for evaluating the procedural accuracy of the achieved plane compared to the preoperative plan, and to assess the achieved results in complex orthopedic procedures, with a main focus on bone tumor resections. The procedural accuracy of the achieved plane will be determined in comparison with the planned plane. To our knowledge, this is the first study to compare different methods for assessing procedural accuracy.

For simplicity, computer-assisted surgery (CAS) and/or PSI will be used as the overarching terms to refer to computer-generated 3D modeling, navigation, and PSIs in this thesis.

2.0 Materials and Methods

2.1. Patients

This single center, retrospective study evaluates the procedural accuracy in a cohort of patients who underwent a complex orthopedic procedure, including correction osteotomies or bone tumor resections, using a 3D planned surgical approach with digital visualization and PSIs (see Table 2.1 and 2.2 for full details). Cutting planes were extracted both in patients and in allografts. In cases involving reconstruction with an allograft, a separate PSI was also designed for the allograft. These planes were included in the analysis but assessed in separate series, and are referred to as "allograft cases". The procedures were performed between 2021 and 2024 at the Leids Universitair Medisch Centrum (LUMC). Surgical margins in case of oncological procedures, were classified as positive (R2), potential microscopic residual disease (R1, tumor at edge of resection specimen), or radical with no residual disease (R0).

Tumor volume (V_{tumor}) was estimated by approximating the tumor as a cylinder. The radius of this cylinder was calculated as the average of half the tumor's anterior-posterior (w) and lateral-medial (l) dimensions. The squared average radius was then multiplied by pi (π) and the cranial-caudal length (h) of the tumor (see Formula 2.1).

$$V_{tumor} = \pi * (\frac{l*w}{4})^2 * h \qquad 2.1$$

Table 2.1: Patient demographics and surgical characteristics

Patient #	Pathology/ Diagnosis	Location	Site	Resection Type	Recon-struction	Assistance Type	Post-operative CT to surgery (days)
1	Osteosarcoma	Dia/meta physis	Femur	Hemi-cortical	Capanna	PSI	5
2	Infected segmental defect	Diaphysis	Femur	Osteotomy	Capanna	PSI	31
3	Ewing sarcoma	Diaphysis	Tibia	Intercalary	Capanna	PSI	88
4	Chondrosarcoma	PI	Pelvic	Hemi-cortical	N/A	PSI	46
5	Chondrosarcoma	Dia/meta physis	Femur	Hemi-cortical	Allograft	PSI	13
6	Malunion	Diaphysis	Tibia	Osteotomy	Plate	PSI	277
7	Osteosarcoma	Diaphysis	Femur	Intercalary	Capanna	PSI	5
8	Chondrosarcoma	PI-IV	Pelvic	Hemi-cortical	N/A	PSI+ navigation	18
9	Chondrosarcoma	PII	Pelvic	Hemi-cortical	Allograft	PSI	1
10	Malunion	Epiphysis	Femur	Osteotomy	Plate	PSI	2

Table 2.2: Tumor characteristics

Patient #	Size (cm)	Volume (cm^3)	Margins
1	7x7.5x13.5	557	R0
3	10.8x12.6x2.8	503	R0
4	8.5x5.2x7.0	248	R0
5	3.8x4.5x8.0	108	R0
7	4.8x1.9x14.3	126	R0
8	5.2x2.8x5.4	68	R1
9	1.8x1.9x2.2	6	R0

2.2. Preoperative planning and simulation

Initially, the surgeon provided the biomechanical engineer with a CT scan of the affected limb. In oncological cases, this also included a delineated tumor in STL format, generated based on the fusion of CT and MRI images using Element software (Brainlab, Germany). All 2D imaging data were imported into image-processing software (Mimics®, Materialise, N.V., Leuven, Belgium), where image segmentation was performed to generate a 3D digital model capturing the detailed bone anatomy. For oncological cases, the segmented tumor contours were integrated into this model.

The segmented data were then exported as STL files into medical CAD software (3-matic®, Materialise N.V.) for further planning. The surgeon and engineer collaboratively determined the surgical approach. In correction osteotomies, cutting planes were established to guide translation and rotational corrections, whereas in oncological procedures, they defined intercalary or hemi-cortical resection margins (Fig. 2.1).

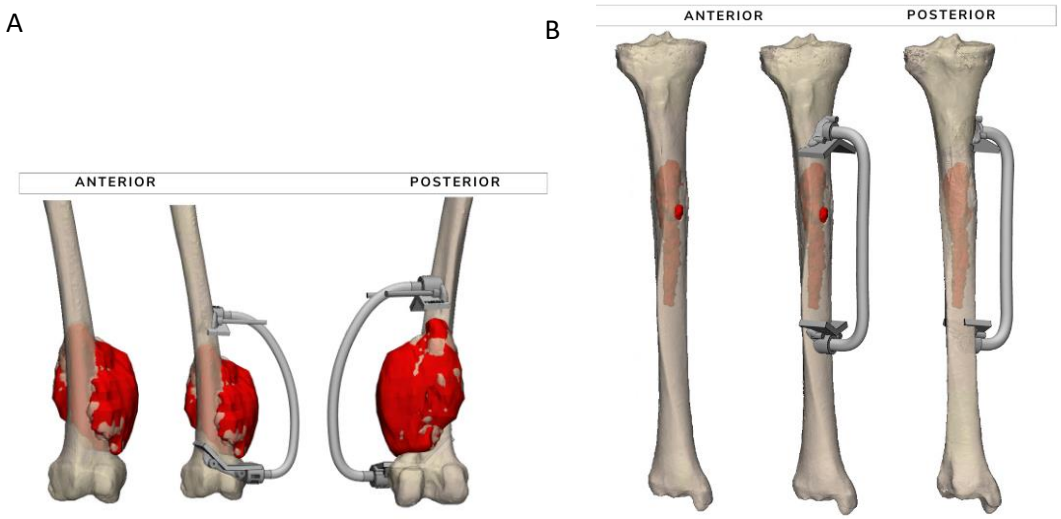


Figure 2.1: Digital 3D models of bone sarcomas including PSIs. (A) Hemi-cortical resection of a osteosarcoma located in the distal femur. (B) Intercalary resection of an Ewing sarcoma located in the midshaft tibia.

Once the preoperative plan was finalized and approved by both the surgeon and the biomechanical engineer, a PSI was designed to align with the intended cutting planes, ensuring intraoperative accuracy. Each PSI was engineered with unique geometries to accommodate the bone morphology of the specific patient (Fig. 2.1). An anatomical model, along with the PSI, was 3D printed in biocompatible polyamide (PA12) using the innovative Selective laser Sintering (SLS) technology by Oceanz (Oceanz B.V., Ede, The Netherlands) and sterilized prior to surgery

2.3. Postoperative accuracy

A scoping review was conducted to identify the most promising methodologies for describing an achieved plane, along with their associated outcome measures, using a subjective scoring system based on grounding criteria (58). Subsequently, an explorative experiment was performed to compare these methodologies in terms of their consistency and ease of use for describing a plane in five cases, which included 18 cutting planes. Finally, in the main experiment, two of the most promising methods were selected. Both the methodologies and their associated outcome measures were compared in ten cases, including 28 cutting planes in patient cases and 17 in allograft cases.

The point cloud method of Ritacco et al. (2013) (27), the mesh portions method of Modabber et al. (2022) (42), and the normal vector method of Brouwer de Koning et al. (2021) (30) showed potential for clearly describing the achieved plane. Additionally, the manual method of Gouin et al. (2014) (23) was tested as it offers a simple and fast approach. Local variations within the achieved plane were analyzed by using the flatness parameter (58). Suitable methods for describing linear and angular outcome measures included the mean distance between points, the mean distance between the center of gravity (COG), the angle between normal vectors, and the pitch and roll angles.

The evaluation of procedural accuracy involved several steps (Fig. 2.2). The planning data consisted of predefined osteotomy planes from the preoperative plan, which in some cases included multiple planes per osteotomy. The actual data were derived from postoperative CT scans. Achieved planes were obtained either through segmentation or by placing points along the osteotomy site in Mimics.

To enable comparison between the planned and achieved planes, 3D digital bone models generated from the segmentation process are required. The postoperative 3D model was first aligned with the preoperative 3D model in 3-Matic. In osteotomy cases, separate alignments were performed for the proximal and distal parts. These alignments used the Iterative Closest Point (ICP) registration technique applied to the corresponding bone models. Subsequently, the achieved planes were extracted from the postoperative 3D models. Finally, the linear and angular deviation between the planned and achieved planes were quantified, including an analysis of local variations within the achieved plane and, for oncological-cases, an assessment of the surgical margin.

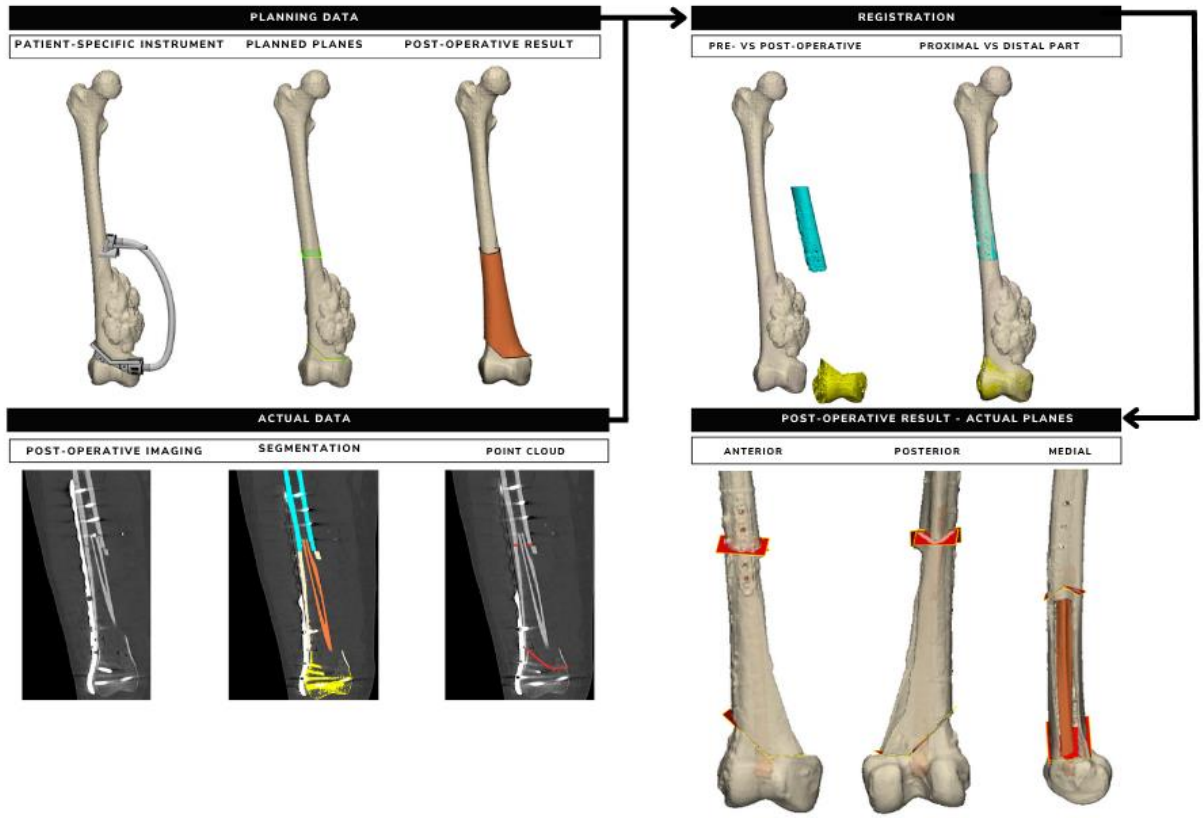


Figure 2.2: The preprocessing steps for data preparation in evaluating procedural accuracy. The planning data, including the patient-specific instrument (PSI), planned planes, and postoperative results, are collected from the preoperative plan, while the actual data is obtained from postoperative imaging. 3D models are created through segmentation, and the achieved plane is determined from the postoperative CT scan. Finally, the achieved 3D models are aligned with the preoperative 3D models, with separate alignments performed for the proximal and distal parts. Planned plane is shown in green, the achieved plane in red, and the allograft in orange.

2.4. Method selection

The achieved planes were determined using different methods. The mesh, normal vector, and manual methods used 3D digital bone models generated from the segmentation process, while the point cloud method used points along the osteotomy site.

2.4.1. Mesh method

For the mesh method areas on the achieved plane in the 3D bone model were manually delineated in 3-Matic until all points on the achieved plane were identified and marked (Fig. 2.3). The marked mesh was then exported as a STL file.

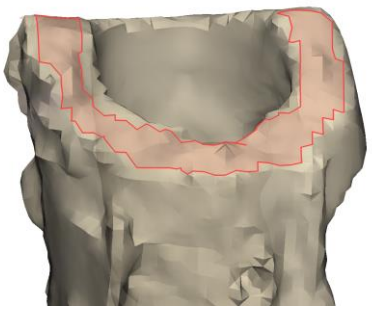


Figure 2.3: Selection of an area on the proximal achieved plane of the 3D bone model by the mesh method

2.4.2. Normal vector method

The normal vector method is another approach to select an area on the achieved plane in the postoperative 3D model in 3-Matic. It uses the plane click mark function with an acceptance angle deviation of five degrees. An initial starting point was manually chosen, and all points within an angular deviation of five degrees between the normal vector of the selected point and the normal vector of the other points in the 3D model were identified as part of the achieved plane (Fig. 2.4A). This process was repeated until all points on the achieved plane were identified and marked (Fig. 2.4B). The marked mesh was then exported as a STL file.

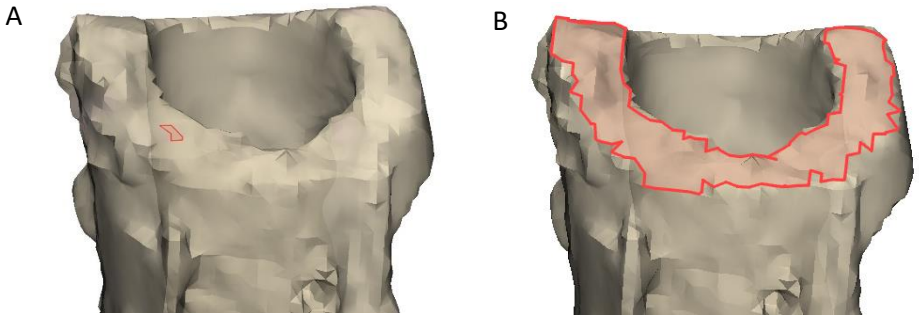


Figure 2.4: Selection of proximal achieved plane on a 3D model using the normal vector method. (A) A single point selected on the achieved plane. (B) Multiple points selected on the achieved plane until all points were marked.

2.4.3. Manual method

The manual method creates a datum plane on the achieved plane by manually defining three points (Fig. 2.5). These points were manually selected such that the resulting datum plane included the achieved plane. The datum plane was then converted to a mesh and exported as a STL file.

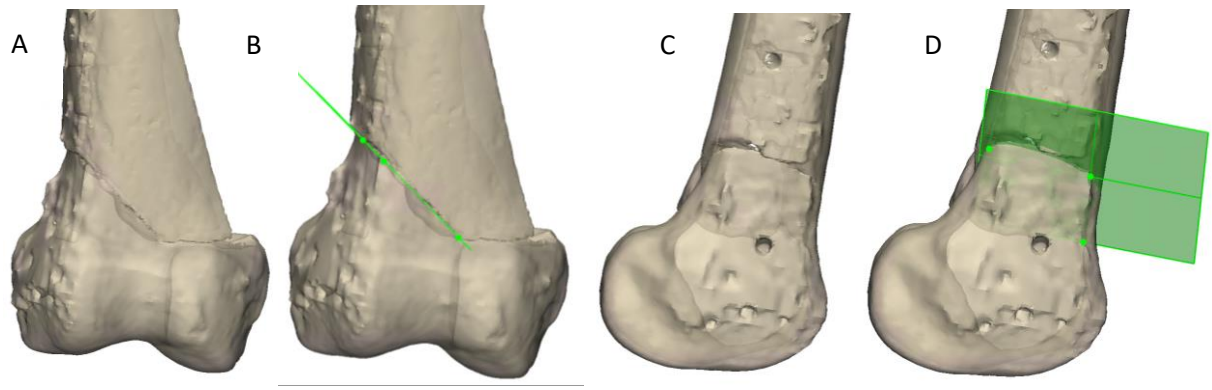


Figure 2.5: Postoperative 3D model of the bone in anterior (A,B) and lateral (C,D) view. (B,D) Selection of three points to determine the distal achieved plane from the bone using the manual method.

2.4.4. Point cloud method

In the point cloud method multiple points on the achieved plane are defined in the CT data directly using Mimics. These points were manually selected on coronal and sagittal slices where the osteotomy was clearly visible. Points were placed across the entire osteotomy plane in that specific slice (Fig. 2.6).

Figure 2.6: Selection of the proximal achieved plane in the postoperative CT scan using the point cloud method.

2.5. Achieved plane

The STL and XML files were imported into the Python code and the points were extracted from these files. The XML files were first converted into CSV format. Subsequently, an initial plane was fitted to the points using singular value decomposition (SVD). The fitted plane was described by the general plane equation (Formula 2.2). A second plane was then fitted to the points within two standard deviations of the initial plane to correct for outliers. This process resulted in the determination of both the planned and achieved planes.

$$z = ax + by + c \qquad 2.2$$

2.6 Local deviations within the achieved plane

The flatness parameter (F) was defined as the minimum distance in millimeters (mm) between two parallel planes that included the achieved plane (Definitions Fig 2.). It was used to quantify the local deviation within the achieved plane. Two methods were used to calculate the flatness.

The first method followed the ISO-1101standard (non-corrected ISO-1101 standard) (59). An initial plane was fitted to the points using SVD. Subsequently, the signed distances from each point to the

plane were calculated, using formula 2.3. The minimum (d_{min}) and maximum distances (d_{max}) were selected to define two parallel planes. The distance between these planes represented the flatness, see Formula 2.4

$$d_i = \frac{ax_i + by_i + cz_i - d}{\sqrt{a^2 + b^2 + c^2}} \quad 2.3$$

$$F = |d_{max} - d_{min}| \quad 2.4$$

The second method (corrected ISO-1101 standard) used the standard deviation. After fitting of the initial plane, the standard deviation of the point-to-plane distances was computed. Two parallel planes were then created, using the points within two standard deviations of the achieved plane. The distance between these planes was used as the flatness measure.

2.7 Outcome measures

Linear and angular outcome measures were used to assess procedural accuracy. Linear deviation is defined as the distance between the planes in mm, while angular deviation measures the difference in angle or orientation between the planes in degrees.

For the angular outcome measures, an additional analysis was performed on the planes that meet the quality criteria. Planes that are highly elongated and/or have a small surface area are more susceptible to inaccurate results. For details on the quality criteria, see Section 2.7.2.1

2.7.1. Linear deviation

For the linear outcome measures, the mean distance to points (Point-to-Point) method and the COG method were used. In the first method, the deviation between the achieved and planned planes was quantified by computing the convex hull (contour) of each plane after projecting the 3D surface points onto the corresponding plane. Within these contours, the mean, minimal, maximal and median signed distances were calculated between the planes. For each point on the grid of the achieved plane, the perpendicular distance to the planned plane was determined, resulting in a local outcome measure (Fig. 2.7).

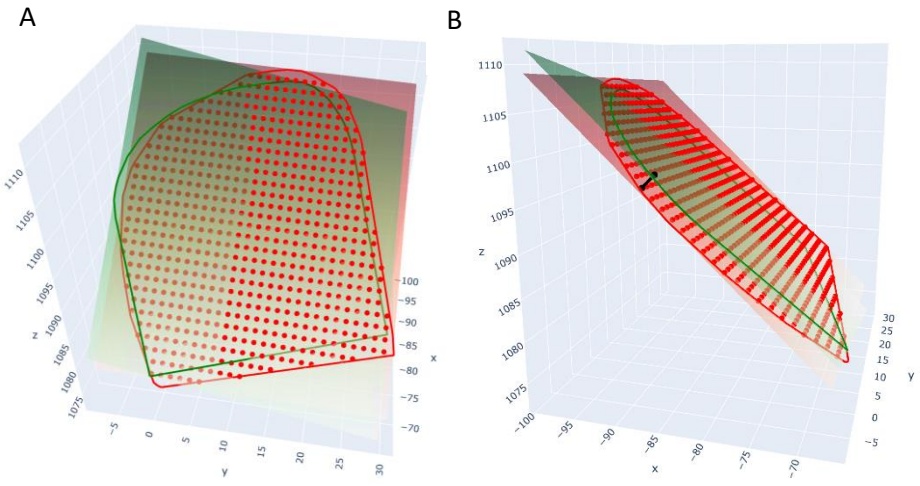


Figure 2.7: Mean distance between points (Point-to-Point) method. The achieved plane is visualized in red and planned plane in green (A) Red points represent the sampled points on the achieved plane. (B) The perpendicular distance (black line) from each point to the planned plane is computed.

The distance between COGs method is focusing on translation error irrespective of direction. The achieved and planned bone contours were resampled at intervals of 1.0 mm and the COG of the contours were calculated. Subsequently, two approaches were used. The first, which is commonly used in the literature, calculated the absolute Euclidean distance between the COG of the achieved plane and that of the planned plane. (Fig. 2.8A). The second approach, determined the COG-to-plane signed distance by averaging the perpendicular distance between the COG of the achieved plane (point) and the planned plane (plane), and vice versa (Fig. 2.8B). Both approaches result in a global outcome measure.

Positive signed distances indicate that less bone was resected relative to the preoperative plan, resulting in closer proximity to the tumor in oncological cases.

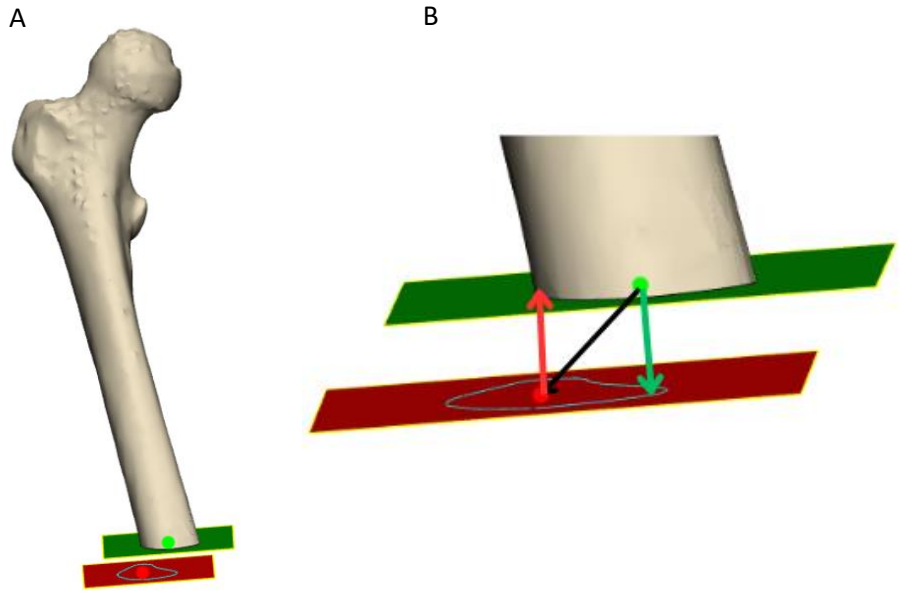


Figure 2.8: Mean distance between COG-based methods. (A) The achieved plane and its COG are represented in red, while the planned plane and its COG are represented in green. (B) COG-COG method: the Euclidian distance between the achieved and planned COGs is visualized in black. COG-plane method: the perpendicular distance from the planned COG to the achieved plane is visualized in green, while the perpendicular distance from the achieved COG to the planned plane is visualized in red.

2.7.2 Angular deviation

For the angular outcome measures, the normal vector method and the roll and pitch method were used.

The normal vector method provides a global measure of orientation error between the achieved and planned planes. First, the normal vectors of the achieved (n_1) and planned (n_2) planes were determined. The unsigned angle (θ) between these vectors was determined by calculating the inner product of the two vectors and dividing it by the magnitude of each vector, using Formulas 2.5 and 2.6. The resulting angle was converted to degrees.

$$\cos(\theta) = \frac{n_1 \cdot n_2}{\|n_1\| \cdot \|n_2\|} \tag{2.5}$$

$$\theta = \arccos\left(\cos(\theta)\right) \tag{2.6}$$

The roll and pitch method is a local, direction-specific metric that captures deviations along clinically relevant anatomical axes. First, the entry and exit points were identified along the resampled contour

on the planned plane. The entry point was defined as the point on the bone contour where the saw blade first contacted the bone. The exit point was determined as the point located along the same axis as the entry point, but on the other side of the bone contour (Fig. 2.9).

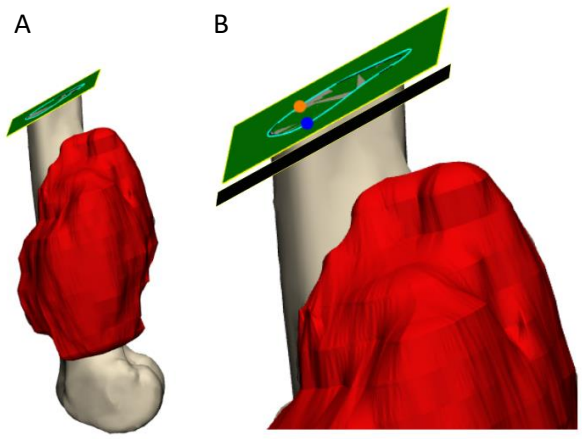


Figure 2.9: Roll and Pitch method (medial view of femur) (A) The planned plane (green), and corresponding bone contour (cyan). (B) Entry point (blue): where the saw blade (black) first contacts the bone. Exit point (orange): directly opposite on the bone contour.

To compute roll and pitch angles, a local right-handed coordinate system was established at the entry point of the planned plane. Three orthogonal axes were defined: (1) the vector from the entry to exit point (x-axis), (2) the normal vector of the planned plane (z-axis), and (3) the cross product of vectors (1) and (2) (y-axis) (Fig. 2.10A). Roll was defined as the rotation of the achieved plane around the y-axis and pitch as the rotation around the x-axis. The roll and pitch angles were then calculated using the arctangent of the cross-product and inner product between the local planned and achieved normal vectors, thereby preserving directional information regarding the rotational deviations. A positive roll was defined as a clockwise rotation of the achieved plane relative to the planned plane (Fig. 2.10B), while a positive pitch was defined as the achieved plane being tilted upwards (Fig. 2.10C).

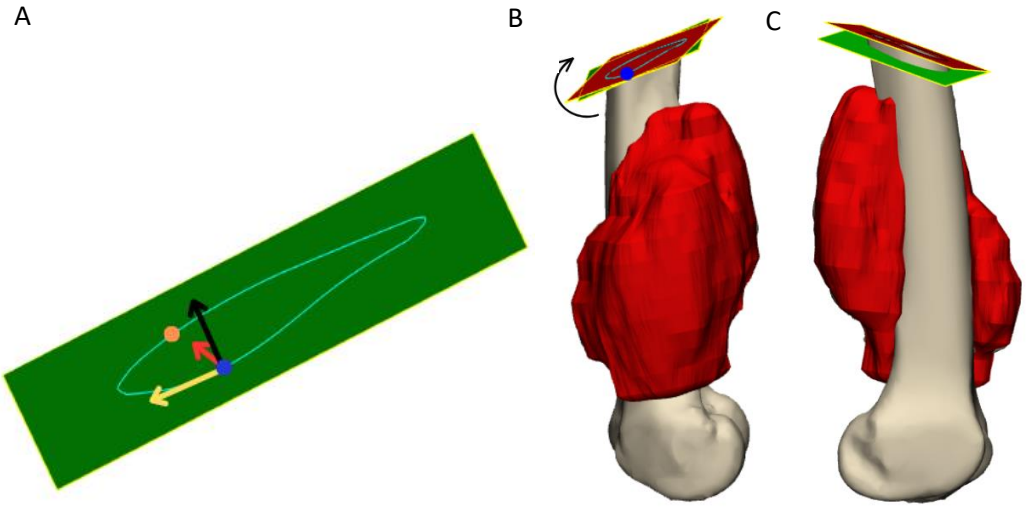


Figure 2.10: Roll and Pitch method. (A) Local coordinate system at entry point (blue) on planned plane (green). X-axis (red): vector from entry to exit (orange) point; Z-axis (black): normal vector of planned plane; Y-axis (yellow): cross product of the x- and z-axes. (B) Medial view of the femur, with the planned plane, and the achieved plane (red). A positive roll is defined as a clockwise rotation of the achieved plane relative to the planned plane. (C) Lateral view of the femur. A positive pitch is defined as an upward tilt of the achieved plane.

2.7.2.1. Quality criteria

To select planes that meet the quality criteria, the general size and shape of the point cloud was determined by conducting a principal component analysis (PCA) on the points of the planned plane ($P_{planned}$) and the filtered points of the achieved plane ($P_{filtered}$), which included all points within the two standard deviations from the initially fitted plane. The COGs of the planes were determined using the corresponding resampled contour points. Subsequently, the centered point sets ($C_{planned}$ and $C_{achieved}$) were constructed by substracting the respective COG from each point, as described by formula 2.7. Covariance matrices for the planned ($A_{planned}$) and achieved ($A_{achieved}$) planes were computed using formula 2.8. The eigenvalues ($\lambda_{planned}$, $\lambda_{achieved}$) were then obtained by solving formula 2.9, involving the identity matrix (I).

$$C_{planned} = \{P_{1planned} - COG_{planned}, \dots, P_{Np\,planned} - COG_{planned}\} \quad (2.7)$$

$$= \{C_{1planned}, \dots, C_{Np\,planned}\}$$

$$C_{achieved} = \{P_{1filtered} - COG_{achieved}, \dots, P_{Npfiltered} - COG_{achieved}\}$$

$$= \{C_{1achieved}, \dots, C_{Np\,achieved}\}$$

$$A_{planned} = C_{planned} {C_{planned}}^{T} \quad (2.8)$$

$$A_{achieved} = C_{achieved} {C_{achieved}}^{T}$$

$$det(\sum - \lambda_{planned} I) = 0 \quad (2.9)$$

$$det(\sum - \lambda_{actual} I) = 0$$

The resulting eigenvalues describe the spread of the points along the head-axis. The smallest eigenvalue corresponds to the shortest axis of an ellipse (radius), while the largest eigenvalue corresponds to the longest axis (radius) (Fig. 2.11). Planes with a smallest eigenvalue below 10.0 mm were considered highly elongated, potentially making the angular outcomes unreliable.

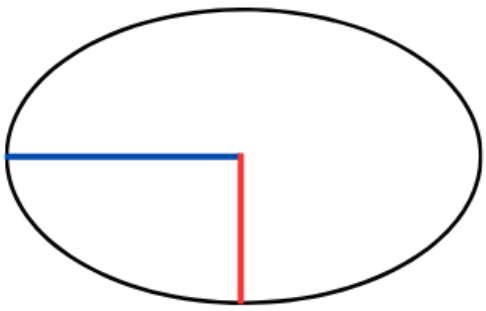


Figure 2.11: Ellipse. The smallest eigenvalue corresponds to the shortest axis (red). Largest eigenvalue corresponds to the longest axis (blue).

2.8. Surgical margin

The surgical margin (SM) was defined as the signed minimum distance from the osteotomy plane to the tumor, as described by formula 2.10 (Definitions Fig. 1). To calculate this, the closest point on the tumor surface (p_{tumor}) to the osteotomy plane was identified. The surgical margin was then computed as the signed distance from this tumor point to the corresponding perpendicular point on the osteotomy plane (p_{plane}), ensuring that this point lay within the anatomical bone contour.

$$SM = \left(p_{tumor} - p_{plane}\right) \tag{2.10}$$

2.9 Statistics

Descriptive statistics, such as frequencies and percentages for categorical variables, were used to describe the study population. Deviations between achieved and planned planes were described using the mean (± SD), minima, maxima, and median. Statistical differences between the mean values of flatness, linear and angular outcome measures, and surgical margin obtained with the mesh and point cloud methods were assessed using a two-sided paired t-test, with p-values <0.05 considered statistically significant.

An overall repeated measures ANOVA test was performed to assess differences among the three approaches for quantifying linear deviation, as well as differences in linear and angular deviations and flatness values across the four methods. Sphericity was checked with Mauchly's test; when this was statistically significant, the Greenhouse-Geisser correction was applied when epsilon < 0.75 or the Huynh-Feldt correction when epsilon > 0.75. If the corrected p-value was significant, pairwise comparisons were performed using post-hoc LSD corrected tests, reporting mean differences with 95% confidence intervals (CIs).

3.0 Results

3.1 Study characteristics

This study included ten patients, with a total of 28 cutting planes in patients and 17 cutting planes in allografts. Three cutting planes were excluded from analysis: one in a patient case and two in allograft cases.

The most common indication was orthopedic oncology, with 4 patients (40%) diagnosed with chondrosarcoma, 2 patients (20%) with osteosarcoma, and 1 patient (10%) with Ewing sarcoma. Clear surgical margins were achieved in all patients, with 6 patients (86%) demonstrating wide negative margins (R0), and 1 patient (14%) exhibiting close negative margins (R1). Tumor volumes ranged from 6 to 557 cm^3, with a mean volume of 231 ± 201 cm^3 (Table 2.1 and 2.2)

Among the non-oncological cases, 2 patients (20%) were diagnosed with malunion, and 1 patient (10%) with an infected segmental defect. The majority of bone defects were located in the diaphysis (4 patients, 40%), followed by the diaphyseal/metaphyseal junction (2 patients, 20%) and the epiphysis (1 patient, 10%). In the remaining three cases, the defect was located in the pelvis. One defect (10%) was located in pelvic region PI according to the Enneking classification, another one (10%) in regions PI-IV, and the last one (10%) in region PII. In terms of anatomical location, most defects were located

in the femur (5 patients, 50%), followed by the pelvis (3 patients, 30%) and the tibia (2 patients, 20%) (Table 2.1).

Resection types included 5 hemi-cortical resections (50%) and 2 intercalary resections (20%). The remaining 3 patients underwent surgery for a correction osteotomy. Hemi-cortical resections were most frequently performed for chondrosarcoma (4 patients, 40%) and once for osteosarcoma (1 patient, 10%). Intercalary resections were performed in cases of Ewing sarcoma (1 patient, 10%) and osteosarcoma (1 patient, 10%) (Table 2.1).

Regarding reconstruction methods, 4 patients (40%) received a combination of a vascularized fibula and allograft (Capanna), 2 patients (20%) received allografts alone, 2 patients (20%) were treated with plate fixation only, and the remaining 2 patients (20%) did not need a reconstruction. PSIs were used in all patients (100%), and navigation was additionally used in 1 patient (10%). The postoperative CT scan was performed on average 48.6 ± 80.3 days after surgery (Table 2.1)

3.2. Method selection

3.2.1 Achieved plane

The mesh, normal vector, and manual methods require a segmented postoperative 3D model, whereas the point cloud method requires a selection of points on a 2D CT scan along the contour of the removed bone.

3.2.2. Linear deviation

Linear deviation between two planes did not show statistically significant differences across the four methods when assessed using the Point-to-Point and COG-to-plane approaches, as indicated by repeated measures ANOVA ($p=0.13$, $p=0.79$, respectively; Table A1). However, a statistically significant difference was observed within the COG-to-COG-approach ($p = 0.01$; Table A1). Post-hoc pairwise comparisons showed statistically significant difference between the manual and mesh methods (3.23 mm 95% CI [1.21,5.20], $p=0.003$), between the manual and normal vector methods (2.58 mm 95% CI [0.38,4.78], $p=0.02$), and between the mesh and point cloud methods (-1.91 mm 95% CI [-3.39,-0.44], $p=0.01$).

The numbers of points used for plane fitting varied considerably between methods, with the mesh and normal vector methods using substantially more points (2552.1 ± 1875.1 and 2054.1 ± 1276.2, respectively; Table A1) than the manual and point cloud methods (6.0 ± 6.0 and 23.4 ± 23.1, respectively; Table A1).

3.2.3. Angular deviation

Angular deviation between two planes did not show statistically significant differences across the four methods when assessed using the angle between normal vector and roll approach, as indicated by repeated measures ANOVA ($p=0.38$, $p=0.40$, respectively; Table A1). However, a statistically significant difference was observed within the pitch approach ($p = 0.03$; Table A1). Post-hoc pairwise comparisons showed statistically significant difference between the manual and normal vector methods (3.88 mm 95% CI[0.34,7.42]), and between the normal vector and point cloud methods (-4.13 mm 95% CI[-7.55,-0.70]).

3.2.4. Flatness

Flatness values obtained with both corrected and non-corrected ISO-1101 standards differed significantly across the four methods, as indicated by repeated measures ANOVA ($p<0.001$; Table A2). Within the non-corrected ISO-1101 standard, post-hoc pairwise comparisons showed significant differences between the manual and mesh methods (-3.92 mm 95% CI[-4.75,-3.08]), the manual and normal vector methods (-4.05 mm 95% CI[-5.55,-2.56]), and the manual and point cloud methods (-3.70 mm 95% CI[-5.35,-2.05]). Similar results were observed for the corrected ISO-1101 standard.

3.3 Outcome measures

3.3.1 Linear outcomes

Based on paired t-tests, no statistically significant differences were observed between the mesh and point cloud methods for either the Point-to-Point or the two COG-based approaches. In the Point-to-Point approach the mean difference between the mesh and point cloud methods was 0.26 mm (95% CI [-0.15, 0.66], $p=0.21$) in patient cases and -0.70 mm (95% CI [-1.54,0.14], $p=0.10$) in allograft cases. Similarly, in the COG-based approaches, all mean differences had p-values > 0.05 (see Figure 3.1 and in Appendix Tables B1 and B2).

In contrast, the overall repeated measures ANOVA revealed statistically significant differences between the three linear quantification approaches (Point-to-Point, COG-to-plane, and COG-to-COG) for both mesh and point cloud methods ($p < 0.001$). In patient cases, post-hoc analysis showed no significant difference between the Point-to-Point and COG-to-plane approaches for either the mesh (-0.53 mm, 95% CI [-1.37, 0.31], $p=0.21$) or the point cloud method (-0.81 mm, 95% CI [-1.59, -0.20], $p=0.05$). However, the COG-to-COG approach differed significantly from both the Point-to-Point (mesh: 3.62 mm, 95% CI [2.32,4.91], $p<0.001$; point cloud: 5.01 mm, 95% CI [3.92, 6.09], $p<0.001$) and the COG-to-plane approach (mesh: 3.08 mm, 95% CI [2.35,3.82], $p<0.001$; point cloud: 4.20 mm, 95% CI [3.27,5.13], $p<0.001$). Similar results were found in allograft cases (see in Appendix Tables B1 and B2).

The procedural accuracy of the point cloud method was -0.77 ± 2.03 mm (Point-to-Point), 0.04 ± 0.18 mm (COG-to-plane), and 4.24 ± 2.39 (COG-to-COG) in patient cases. In allograft cases, these values were 0.47 ± 2.32 mm, -0.03 ± 0.15 mm, and 4.41 ± 2.44 mm, respectively. For the mesh method, procedural accuracy was: -0.51 ± 2.16 mm, 0.02 ± 0.19 mm, and 3.10 ± 1.89 mm, respectively in patient cases, and -0.23 ± 1.60 mm, -0.07 ± 0.26 mm, and 4.22 ± 2.27 mm, respectively in allograft cases (see Figure 3.1 and in Appendix B Tables B1 and B2). For the Point-to-Point approach, the procedural accuracy is also reported separately for each patient, see Tables B3-B4 in Appendix B.

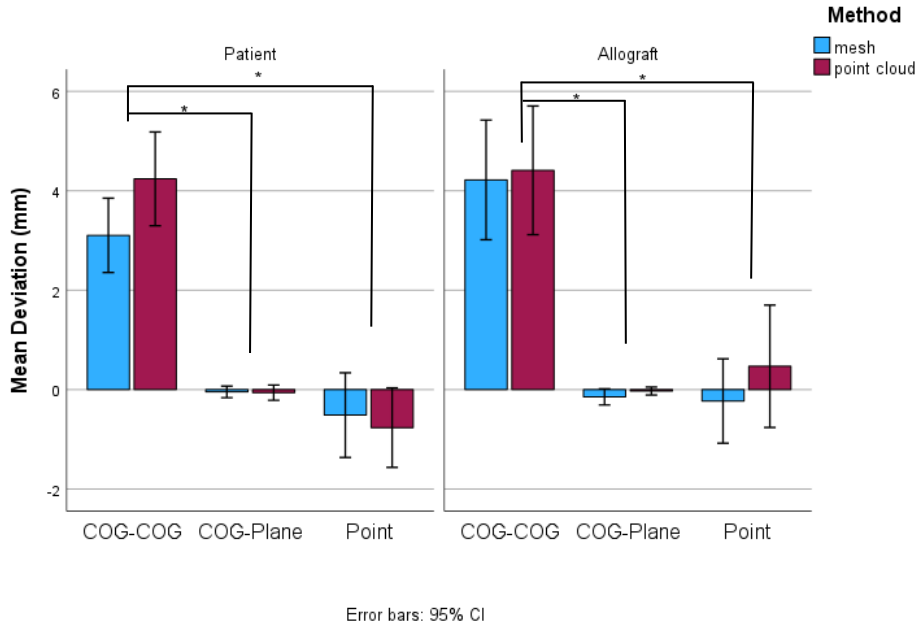


Figure 3.1: Mean linear deviation in mm between the planned and achieved planes for the COG-to-COG approach, COG-to-Plane approach and Point-to-Point approach in patient and allograft cases. The mesh method is presented in blue and the point cloud method in red. The values are positive(+)/negative(-) representing less bone resected compared to the planning. Statistical significant difference is indicated with (*). Patients: N=28; Allografts: N=17

3.3.2 Angular outcomes

Paired t-test showed no statistically significant differences between the mesh and point cloud methods for the angle between normal vectors (patients: 0.47 degrees (95% CI [-1.07,2.01], p=0.53; allografts: 0.62 degrees (95% CI [-1.78,3.02], p=0.59), for roll angle (patients: -0.07 degrees (95% CI [-1.12,0.98], p=0.84; allografts: -0.67 degrees (95% CI [-1.67,0.33], p=0.17), and for pitch angle (patients: 0.79 degrees (95% [-1.64,3.22], p=0.51; allografts: -1.96 degrees (95% CI [-4.63,0.70], p=0.14), (see Figure 3.2 and in Appendix Tables C1 and C2).

The procedural accuracy of the point cloud method was 6.13 ± 5.43 degrees (normal vector), 0.15 ± 5.36 degrees (pitch), and -0.21 ± 3.88 degrees (roll), in patient cases. Values of 4.71 ± 2.92 degrees, 2.91 ± 3.85 degrees, and 0.70 ± 2.30 degrees, respectively, were observed in allograft cases. For the mesh method, procedural accuracy was 6.60 ± 6.50 degrees, -0.63 ± 6.84 degrees, and -0.14 ± 4.01 degrees, respectively in patient cases. Values of 5.33 ± 5.07 degrees, 0.98 ± 7.08 degrees, and 0.03 ± 1.60 degrees, respectively, were observed in allograft cases (see Figure 3.2 and in Appendix Tables C1 and C2). For the normal vector approach, the procedural accuracy is also reported separately for each patient, see Tables C3-C4 in Appendix C.

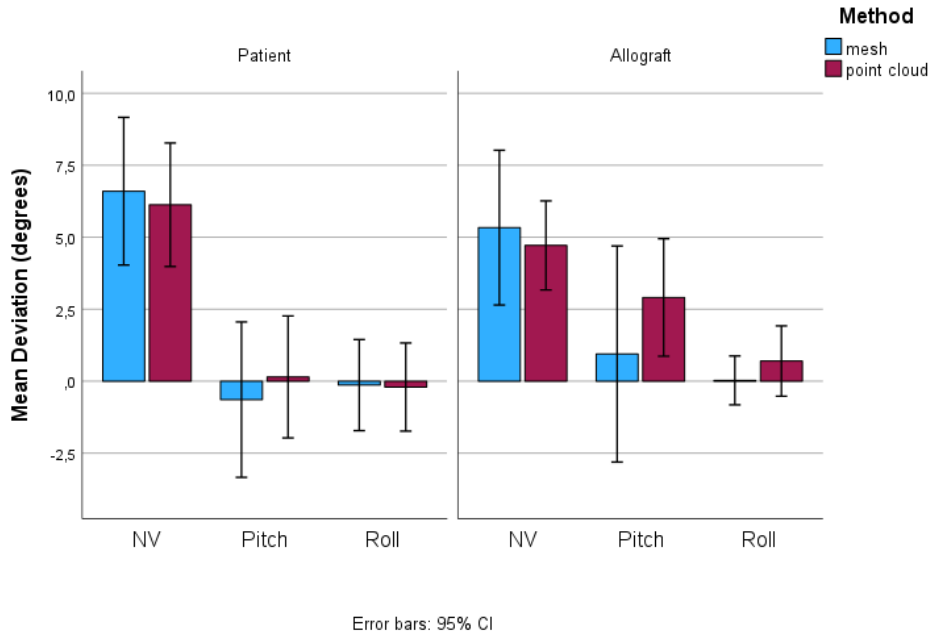


Figure 3.2: Mean angular deviation in degrees between the planned and achieved planes for the normal vector (NV) approach, and pitch and roll approach in patient and allograft cases. The mesh method is presented in blue and the point cloud method in red. The Pitch and Roll values are positive(+)/negative(-) representing the achieved plane positioned upwards and rotated clockwise compared to the planning. Patients: N=28; Allografts: N=17.

3.3.2.1 Quality assessment

In the quality assessment, a subset of planes was selected for further analysis. This resulted in the selection of 24 planes in patient cases and 12 in allograft cases using the mesh method, and 20 and 14 planes, respectively, using the point cloud method. See appendix D for the scatterplots visualizing the selected planes in patient cases (Fig. D1A) and allograft cases (Fig. D1B) for the mesh method.

Based on the normal vector approach, a mean angular deviation of 5.64 ± 5.32 degrees was observed in patient cases using the mesh method and 5.02 ± 2.68 degrees using the point cloud method. In allograft cases, the mean angular deviation was 3.75 ± 2.11 degrees using the mesh method and 4.98 ± 2.95 degrees using the point cloud method.

Based on the pitch and roll angle approach, mean values of 1.39 ± 3.76 degrees and -0.39 ± 1.20 degrees were observed in patient cases using the mesh method, whereas values of 2.85 ± 4.08 degrees and 0.66 ± 2.52 degrees were observed using the point cloud method. In allograft cases, mean values of -1.39 ± 3.76 degrees and -0.39 ± 1.20 degrees were observed using the mesh method, whereas values of 2.85 ± 4.08 degrees and 0.66 ± 2.52 degrees were observed using the point cloud method. See Appendix D Figure D2 for the boxplots illustrating the mean angular deviation before and after quality assessment.

3.3.3 Flatness

Paired t-test showed no significant difference between the mesh and point cloud methods for the corrected ISO-1101 standard in patient cases (0.30 mm 95% CI [-0.54,1.15], p=0.47). However, a significant difference was found in allograft cases (0.90 mm 95% CI [0.35, 1.46], p=0.003). For the non-corrected ISO-1101 standard, significant differences were observed in both patient cases (1.70 mm 95% CI [0.71,2.68], p=0.002) and allograft cases (1.67 mm 95% CI [0.99,2.35], p< 0.001) (Figure 3.3, and Appendix Tables E1 and E2).

For both patient and allograft cases, a statistical significant difference was observed between the corrected and non-corrected ISO-1101 standards when the mesh method was used (patient: -1.55 mm 95% CI [-1.86, -1.23],p < 0.001; allograft: -0.90 mm 95% CI [-1.24,-0.56], p < 0.001). However, no significant difference was found when the point cloud method was used (patient: -0.15 mm 95% CI [-0.38,0.07],p = 0.17; allograft: -0.14 mm 95% CI [-0.28,0.01], p= 0.07) (Figure 3.3, and Appendix Tables E1 and E2).

Mean flatness values were 2.78 ± 1.19 mm (mesh) and 2.48 ± 2.14 mm (point cloud) using the corrected ISO-1101 standard in patient cases. According to the non-corrected standard, mean flatness values were 4.33 ± 1.74 mm (mesh) and 2.63 ± 2.30 mm (point cloud). In allograft cases, mean flatness values were 2.36 ± 0.90 mm (mesh) and 1.46 ± 0.62 mm (point cloud), according to the corrected ISO-1101 standard, and 3.29 ± 1.17 mm and 1.60 ± 0.69 mm, respectively, according to the non-corrected standard (see in Appendix Tables E1 and E2).

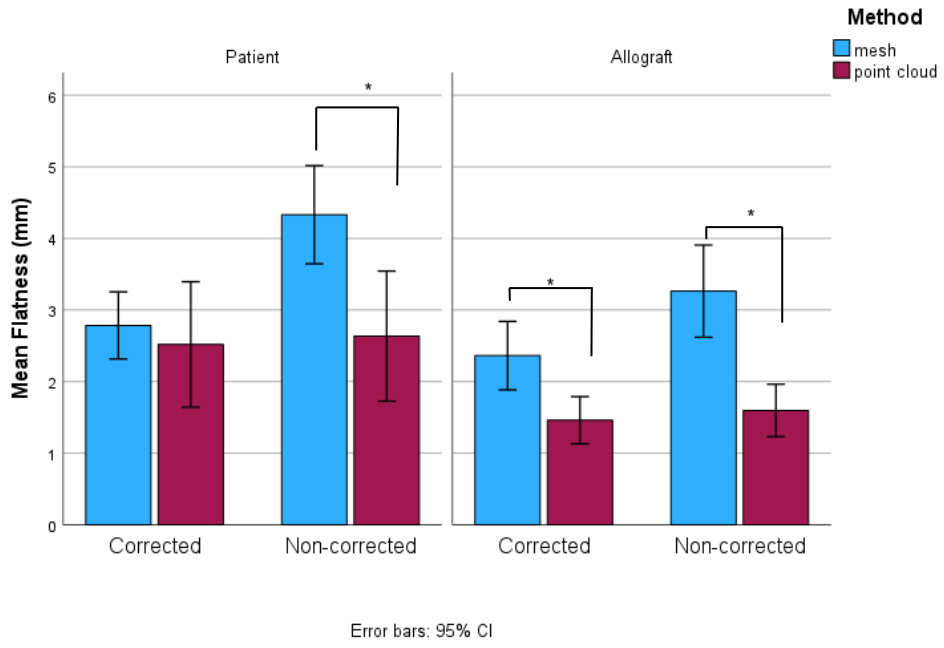


Figure 3.3: Mean Flatness in mm of the achieved planes for the corrected and non-corrected ISO-1101 standards. The mesh method is presented in blue and the point cloud method in red. Statistical significant difference is indicated with (*). Patients N=28; Allografts: 17.

3.3.4 Surgical margin

Paired t-test showed significant difference between the mesh and point cloud methods for the surgical margin (0.9 mm 98% CI [0.24,1.56], p=0.01) (Figure 3.4, and Table F1 in Appendix F). The achieved surgical margin measured with the mesh method was 10.2 ± 5.8 mm, and 11.1 ± 5.7 mm for the point cloud method (see Table F1 in Appendix F)

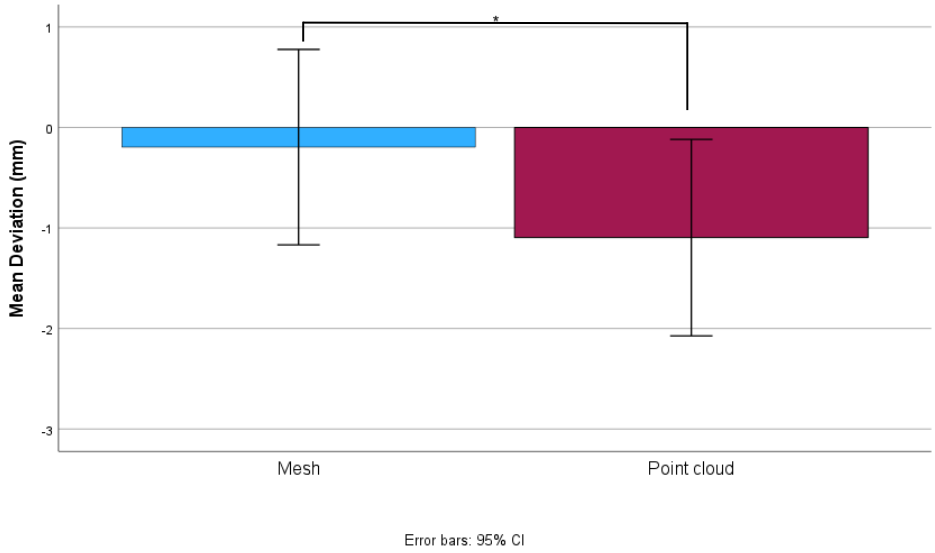


Figure 3.4: Mean deviation (mm) in surgical margin between planned and achieved planes for the mesh and point cloud methods. The mesh method is presented in blue and the point cloud method in red. Negative values(-) indicate that the achieved plane was cut farther from the tumor than planned. Statistical significant difference is indicated with (*). N=23.

4.0 Discussion

Preoperative 3D planning has become the standard of care for complex orthopedic procedures, including multiplanar osteotomies and tumor resections. Many studies have reported the procedural accuracy between the 3D planned and achieved planes (6, 11, 19, 23, 26-57). However, as standardized measurement for procedural accuracy is still lacking, the studies are hard to compare. Standardized measurements would enable more reliable comparisons across studies and contribute to the development of best practices.

This study therefore aimed to investigate the best suitable standardized approach for evaluating the procedural accuracy of the achieved planes compared to the preoperative plane, and to assess the achieved results in complex orthopedic procedures. Although only PSIs were used in this study, the findings are also applicable to CAS. The in-house assessment of the different methods described in literature clearly demonstrated the need for a standardized validated methodology.

4.1 Results

4.1.1Method selection

The results highlight clear differences in usability among the four evaluated techniques (manual, normal vector, mesh and point cloud methods) for achieved plane determination.

The manual method requires a complete 3D bone model, which can be generated relatively quickly. However, this method is highly subjective, relying heavily on user input and interpretation (Fig. 2.5). The normal vector method requires separate 3D models of the different bone segments (e.g. proximal part, distal part, resected part), making it more time-consuming. Additionally, this method is only suitable for relatively flat planes, as it becomes highly time-consuming for irregular surfaces (Fig. 2.4). The mesh method also depends on separate 3D bone models, but offers a more objective and reproducible approach compared to the manual and normal vector methods, and is faster than the normal vector method. The point cloud method allows for quick and simple positioning of points on the achieved plane and is the only method that does not require segmentation, which is an advantage as not everyone has the expertise to perform such a segmentation. In summary, both the mesh and point cloud methods are the best in terms of ease of use and objectivity compared to the manual and normal vector methods, with the point cloud method being the most accessible due to the lack of segmentation requirements.

Furthermore, analyses of angular deviations and the flatness parameter within a subset of the data, demonstrated that the point cloud method performed comparably to the mesh method, with no statistically significant differences observed across the evaluated approaches ($p>0.05$). In contrast, analyses of linear deviations in this subset revealed a statistically significant difference between the mesh and point cloud methods within the COG-to-COG approach (-1.91 mm 95% CI [-3.39,-0.44], $p=0.01$). No statistically significant differences were found in the other two linear deviation approaches. Since the COG-to-COG approach is not considered a reliable measure for linear quantification (see below), this result does not affect the overall conclusion that the point cloud method performed comparably to the mesh method. This finding is promising for future research, particularly when larger datasets are assessed to quantify the linear and angular deviations between the planes.

4.1.2 Linear outcomes

The linear results, measured across the entire datatset, indicate that there is no significant difference in mean distance between the point cloud and mesh methods for all three linear deviation approaches, in both patient and allograft procedures ($p>0.05$). However, when comparing the linear deviation approaches themselves, the COG-to-COG approach differed significantly from both the COG-to-plane and Point-to-Point approaches. For patients assessed with the mesh method, the COG-to-COG approach differed from the Point-to-Point approach by 3.62 mm (95% CI [2.32,4.91], $p < 0.001$) and from the COG-to-Plane approach by 3.08 mm (95% CI [2.35,3.82], $p < 0.001$). Using the point cloud method, COG-to-COG differed from Point-to-Point by 5.01 mm (95% CI [3.92,6.09], $p < 0.001$) and from COG-to-Plane by 4.20 mm (95% CI [3.27,5.13], $p < 0.001$)). Similar results were observed in allograft cases. The COG-to-COG approach is more sensitive to errors introduced by plane translation or inaccuracies in the convex hull. This is illustrated in Figure G1 in Appendix G, where the COG-to-COG approach yielded a deviation of 10.23 mm, while the COG-to-plane approach showed a deviation of only -0.16 mm. Despite these disadvantages, all studies included in the scoping review that used a COG-based method, applied the COG-to-COG approach (30, 36, 40, 51), suggesting that previously reported deviations may have been overestimated.

Previous studies reported much smaller differences (approximately 0.5 mm) between COG-to-COG and Point-to-Point approaches. For example, Brouwer de Koning et al. (2021) reported a mean deviation of 0.9 ± 0.5 mm for the anterior plane and 2.0 ± 1.0 mm for the posterior plane in the maxillofacial region, using the COG-to-COG approach (30), while Liu et al. (2014) reported a mean deviation of 1.39 ± 1.05 mm in the same region using the Point-to-Point approach (41). This discrepancy may be attributed to anatomical differences. The relatively small planes in the maxillofacial region reduce the sensitivity of the COG-to-COG approach to plane translation and convex hull inaccuracies.

A procedural accuracy of -0.77 ± 2.03 mm was observed between the planned and achieved planes using the point cloud method and Point-to-Point approach. This implies that the achieved plane was positioned further from the planned plane, resulting in more bone resected than intended in the preoperative plan. In contrast, the study of Ter Braak et al. (2020) reported an accuracy of 1.1 ± 0.6 mm in plaster mandibular models (51). The higher variability observed in our study may reflect the complexity of working with actual bones with potentially unclear planes instead of plaster models

4.1.3 Angular outcomes

Two approaches were used to quantify angular outcome measures. The normal vector approach provides a global measure of orientation error between the achieved and planned planes. This method is computationally simpler and well suited for quick assessments, although it lacks directional specificity. In contrast, the roll and pitch method is a local, direction-specific metric that captures deviations along clinically relevant anatomical axes. However, it is more difficult to implement and interpret, particularly for non-technical users.

Additionally, the results, measured across the entire dataset, indicate that there is no significant difference between the mesh and point cloud method when the normal vector, and roll and pitch approaches were used ($p>0.05$).

For patients, the point cloud method yielded a procedural accuracy of 0.15 ± 5.36 degrees in pitch and -0.21 ± 3.88 degrees in roll between the planned and achieved planes. Following quality assessment, these values changed to 1.21 ± 3.94 degrees and -0.71 ± 2.97 degrees, respectively. Similar results

were observed with the mesh method and in allograft cases. Excluding elongated planes resulted in a more reliable estimation of angular deviation, evidenced by a reduced standard deviation. The observed shift in the mean indicates a possible bias caused by the inclusion of less reliable cutting planes in the first analysis.

The study of Hasan et al. (2020) reported angular deviations of 6.7 ± 4.6 degrees in roll and 5.4 ± 1.5 degrees in pitch in the mandibular region, using a navigation-assisted approach (39). The lower angular deviation observed in the present study may be attributed to differences in surgical guidance technique. Furthermore, anatomical differences likely play a role as the mandible presents smaller and more constrained bone surface areas, which may increase sensitivity to angular inaccuracies, as also reflected in the results of our quality assessment.

4.1.4 Flatness

The results, evaluated across the entire dataset, demonstrate a significant difference between the mesh and point cloud methods for both the corrected and non-corrected ISO-1101 standards in allograft cases (corrected: 0.98 mm 95% CI [0.43,1.54, p=0.002], non-corrected: 1.68 mm 95% CI [1.02,2.35], p<0.001). In patient cases, a significant difference was found only for the non-corrected ISO-1101 standard (1.67 mm 95% CI [0.68,2.65], p=0.002), while the corrected ISO-1101 standard did not show a significant difference (0.44 mm 95% CI [-0.41,1.29], p=0.30). This might be explained by the reduced sensitivity of the corrected standard to outliers. Interestingly, the corrected ISO-1101 standard still revealed a significant difference in allograft cases. This may be attributed to the lower variability observed in the allograft data, resulting in smaller standard deviations and a narrower confidence interval. In this scenario, even a small, consistent difference can become statistically significant. Therefore, future studies with larger samples are needed to establish with greater certainty whether the point cloud method is comparable to the mesh method.

Overall, flatness values were generally 1 mm higher in patient cases than in allograft cases. This may be attributed to the in vivo surgical environment, compared to the more controlled cutting conditions of allograft preparations outside the body.

Furthermore no statistically significant differences were observed within the point cloud method when comparing the corrected and non-corrected standards themselves (patient: -0.15 mm 95% CI [-0.38,0.07], p=0.17; allograft -0.14 mm 95% CI [-0.28,0.01], p=0.07). In contrast, significant differences were found within the mesh method, with an average difference of approximately 1 mm (patient: -1.38 mm 95% CI [-1.69,-1.07], p<0.001; allograft: -0.84 mm 95% CI [-1.15,-0.53], p<0.001). These findings highlight the importance of standardizing the method used to define the achieved plane and show the influence of the chosen standard (corrected vs non-corrected) to the measured outcome.

For patients, the point cloud method yielded a mean flatness of 2.63 ± 2.30 mm using the non-corrected ISO-1101 standard and 2.48 ± 2.14 mm using the corrected standard. For allografts, mean flatness of 1.60 ± 0.69 mm and 1.46 ± 0.62 mm were observed, respectively. The study of Cartiaux et al. (2014) determined flatness using the non-corrected ISO-1101 standard on a synthetic pelvic sawbone model (32). Achieved planes were digitized directly on the 3D bone model using a coordinate measuring machine. They reported an average flatness in the posterior ilium of 0.8 mm (CI 0.6-1.0 mm), which is approximately two times lower than the value measured in our allograft cases (1.60 mm). This discrepancy may be attributed to differences in the type of bones used (synthetic sawbones versus actual bones presenting cortical and trabecular structures).

4.1.5 Surgical margin

Although the error bars in the bar plot overlap (Fig. 3.4), analysis across the entire dataset using a paired-sample t-test revealed a significant difference between the mesh and point cloud method for surgical margin (0.9 mm 95% CI[0.24, 1.56], p=0.01), with the point cloud method showing a higher value. This is because the paired analysis accounts for the within-subject consistency of the differences, which is not reflected in the group-level error bars. The significant difference observed between the methods may be related to their different sensitivities to noise and outliers. The mesh method can be more susceptible to these factors, as noise and outliers can influence the shape of the triangles. In contrast, the point cloud method may be less affected by noise and outliers due to their smaller impact on the overall point distribution. Additionally, the significant difference can also be due to the small sample size (N=23).

A surgical margin difference of -1.1 ± 2.2 mm was observed using the point cloud method. This implies that the achieved plane is generally located farther from the tumor than planned. In future preoperative 3D plannings this may allow for a reduction in the required surgical margins, potentially preserving more healthy tissue without compromising oncological safety.

The study of Evrard et al. (2022) reported a mean surgical margin difference of -0.4 ± 1.8 mm across various bones (6), which differs by 0.7 mm with our result using the point cloud method and by only 0.2 mm when using the mesh method. As the method used to determine the achieved plane in their study is not specified, the observed difference may be attributed to methodological variations.

4.2 Limitations

Several limitations of this study should be acknowledged. First, all measurements were performed by a single observer. The potential impact of inter- and intra-observer variability was therefore not assessed. It is possible that differences between the mesh and point cloud methods could be influenced by subjective interpretation of the cutting planes. Future studies should include repeated assessments by additional observers to validate the robustness of these methods.

Second, registration errors were not corrected when analyzing the linear and angular outcome measures. These errors may have influenced the procedural accuracy results.

Third, cases involving correction osteotomies and bone tumor resections were analyzed together. In oncological cases, the accurate positioning of the PSI is more critical than in non-oncological cases. Therefore, the results of correction osteotomies may have influenced the overall procedural accuracy in oncological cases, as the difference between the planned and achieved planes is likely smaller in these cases.

Fourth, all cutting planes were considered separately. However, cutting planes within the same patient or allograft are often interdependent, as they tend to deviate by a similar factor if the PSI is not perfectly positioned. This may not hold true when a large connection bar is used (Fig. 2.1A). In such cases, the PSI is easier to position distally, while positioning it proximally is more challenging due to diaphyseal localization.

Fifth, the achieved surgical margin can be determined more precisely. In this study, the surgical margin was defined as the minimal distance between the tumor and the achieved plane, without accounting for the kerf (Definitions Fig. 2). Future studies could address this limitation by either subtracting the

saw blade thickness for each individual case or by determining the achieved planes through a postoperative CT scan of the resection specimen, rather than scanning the patient.

Lastly, the limited sample size reduces the reliability of this study. Postoperative CT scans after tumor resections and correction osteotomies are not routinely acquired as part of the clinical protocol. Follow-up often relies on radiographs for correction osteotomies, and radiographs at progressive intervals, with MRI/CT chest at specific intervals, up to 10 years after surgery, for tumor resections. However, postoperative CT scans are crucial for procedural feedback. Therefore a low-dose CT protocol could be implemented to improve this. Additionally, collaboration between multiple hospitals would be beneficial to increase the sample size and enhance the statistical power of future studies.

4.3 Outliers

In total, three cutting planes were excluded from analysis: one in a patient case and two in allograft cases. These cases are show in Appendix H. The excluded patient case involved a PSI with a curved plane designed to match the cartilage. However, according to the orthopedic surgeon, this cutting plane was manually sawn during surgery and therefore deviated from the intended plan. The allograft case concerns a typical Capanna reconstruction, in which the proximal part of the allograft was manually shortened during surgery to adapt to high soft tissue tension. Additionally, a groove was removed from the allograft to accommodate the vascular stem connection. These alterations resulted in cutting planes that were unsuitable for quantitative analysis.

5.0 Conclusion

This study demonstrates that the mesh and point cloud method are well suited for determining the achieved plane in complex orthopedic procedures, particularly in terms of ease of use and objectivity. Analyses of linear and angular deviations showed that the point cloud method performed comparably to the mesh method with an average procedural accuracy of -0.77 ± 2.03 mm, -0.21 ± 3.88 degrees in roll, and 0.15 ± 5.36 degrees in pitch.

Analysis of flatness values indicated that the point cloud method perfomed comparably to the mesh method only when using the corrected ISO-1101 standard, yielding an average flatness of 2.48 ± 2.14 mm. Future research with larger sample sizes is needed to confirm whether the point cloud method is also comparable to the mesh method in allograft cases. Given the advantages of the corrected ISO-1101 standard over the uncorrected standard, the significant difference between these standards, and the potentially non-significant differences observed between the mesh and point cloud method with the corrected standard in allograft cases, future research should carefully consider which approach to employ.

Analysis of surgical margin values indicate that the point cloud method did not perform comparably to the mesh method, yielding an average surgical margin difference of -1.1 ± 2.2 mm and -0.2 ± 2.2 mm, respectively. To exclude the possibility that the observed significant differences are due to the different ways in which the mesh and point cloud methods handle noise and outliers, future studies should include larger sample sizes.

If future studies demonstrate that point cloud method performs comparably to the mesh method across all assessed parameters, the point cloud method should be adopted for determining the achieved plane, as its simplicity and speed of use would result in faster and easier quantification between planned and achieved planes.

Within the linear outcome measures, significant differences were observed between the COG-to-COG approach and other linear approaches for both methods. As the COG-to-COG approach is not considered a reliable outcome measure (see 4.1.2), future research should use alternative linear outcome measures. For global assessment, the COG-to-plane approach is recommended, while the Point-to-Point approach is suitable for local assessment. Regarding angular outcome measures, the angle between normal vectors should be used as global outcome measure, and roll and pitch angles as local measure.

The average outcome measures reported in this study provide a useful baseline for assessing procedural accuracy in complex orthopaedic procedures using 3D planning and CAS. Complementary studies with larger sample sizes are needed to fully validate the results observed here. Once completed, these results could contribute to the development of a more standardized approach for evaluating procedural accuracy in this field.

Appendix

A – Method selection

Table A1: Comparison of four methods to describe the achieved plane, including the number of points used for plane fitting and the resulting outcome measures. N=18. (*) indicates a statistically significant difference.

Planned vs achieved planes - Patient cases		Mesh method	Normal vector method	Manual method	Point cloud method	p-value
Number of points		2552.1 ± 1875.1	2054.1 ± 1276.2	6.0 ± 6.0	23.4 ± 23.1	-
Distance Point-to-Point (in mm)	Mean ± std	-0.28 ± 1.55	0.03 ± 1.64	-0.02 ± 1.22	-0.54 ± 1.54	0.13
	Min	-5.94	-6.01	-7.44	-5.37	
	Max	2.98	4.90	5.13	3.65	
	Median	0.04	0.40	-0.35	-0.71	
Distance COG-to-plane (in mm)	Mean ± std	0.00 ± 0.05	0.04 ± 0.15	0.01 ± 0.35	-0.01 ± 0.16	0.79
	Min	-0.09	-0.09	-0.72	-0.33	
	Max	0.14	0.55	0.78	0.28	
	Median	-0.01	-0.01	0.01	-0.01	
Distance COG-to COG (in mm)	Mean ± std	2.33 ± 1.50	2.95 ± 1.69	5.53 ± 4.22	4.24 ± 2.58	0.01 (*)
	Min	0.79	0.96	0.41	1.35	
	Max	6.57	6.72	16.00	10.22	
	Median	1.99	2.22	4.49	3.24	
Angle between normal vectors (in degrees)	Mean ± std	4.73 ± 2.41	6.32 ± 6.28	5.22 ± 4.37	6.85 ± 5.28	0.38
	Min	1.39	0.30	0.55	2.03	
	Max	10.51	29.43	17.52	25.18	
	Median	4.68	3.96	3.87	5.07	
Angle in roll (in degrees)	Mean ± std	-0.36 ± 2.10	1.63 ± 5.27	-0.07 ± 2.67	-1.25 ± 2.72	0.40
	Min	-5.26	-20.15	-6.70	-7.49	
	Max	4.67	5.00	6.23	4.84	
	Median	-0.26	-0.41	-0.30	-0.43	
Angle in pitch (in degrees)	Mean ± std	1.19 ± 4.46	-2.11 ± 6.52	1.76 ± 5.86	2.01 ± 7.63	0.03 (*)
	Min	-10.31	-21.95	-13.01	-9.29	
	Max	7.79	6.04	17.33	24.27	
	Median	2.08	-0.37	1.89	2.16	

Table A2: Comparison of four methods to describe the achieved plane and resulting flatness outcome. N=18. (*) indicates a statistically significant difference.

Planned vs achieved planes - Patient cases		Mesh method	Normal vector method	Manual method	Point cloud method	p-value
Corrected ISO-1101 standard	Mean ± std	2.62 ± 1.11	2.79 ± 2.10	0.00 ± 0.01	3.49 ± 3.17	<0.001 (*)
	Min	1.01	0.48	0.00	0.47	
	Max	4.71	10.38	0.03	11.75	
	Median	2.41	2.04	0.00	2.17	
Non corrected ISO-1101 standard	Mean ± std	3.92 ± 1.63	4.06 ± 2.92	0.01 ± 0.02	3.70 ± 3.22	<0.001 (*)
	Min	1.48	1.49	0.00	0.44	
	Max	7.48	15.20	0.08	11.28	
	Median	3.24	3.56	0.00	2.69	

B – Linear outcomes

Table B1: Comparison of the mesh and point cloud methods across three quantification approaches, and resulting linear outcome measures in patient cases. The lower part presents pairwise comparisons between the approaches within both the mesh and point cloud methods, including mean differences, 95 CIs, and p-values. (*) indicates a statistically significant difference. N=28.*

Planned vs achieved planes Patient		**Mesh method**	**Point cloud method**	**Mean difference (95% CI)**	**p-value**
Number of points		2771.9 ± 1964.5	27.8 ± 25.8	-	-
Distance Point-to-Point (in mm)	Mean ± std	-0.51 ± 2.16	-0.77 ± 2.03	0.26 [-0.15, 0.66]	0.21
	Min	-7.60	-5.37		
	Max	5.32	6.82		
	Median	-0.21	-1.27		
Distance COG-to-plane (in mm)	Mean ± std	0.02 ± 0.19	0.04 ± 0.18	-0.02 [-0.10,0.06]	0.62
	Min	-0.66	-0.27		
	Max	0.59	0.54		
	Median	0.00	0.02		
Distance COG-to COG (in mm)	Mean ± std	3.10 ± 1.89	4.24 ± 2.39	-1.14 [-2.30,0.03]	0.06
	Min	0.79	1.35		
	Max	8.79	10.22		
	Median	2.42	3.39		
Pairwise comparisons					
Point-to-Point vs COG-to-plane		-0.53 [-1.37, 0.31], p=0.21	-0.81 [-1.59,-0.20], p=0.05		
COG-to-COG vs Point-to-Point		3.62 [2.32,4.91], p<0.001 (*)	5.01 [3.92,6.09], p<0.001 (*)		
COG-to-COG vs COG-to-plane		3.08 [2.35,3.82], p<0.001 (*)	4.20 [3.27,5.13], p<0.001 (*)		

Table B2: Comparison of the mesh and point cloud methods across three quantification approaches, and resulting linear outcome measures in allograft cases. The lower part presents pairwise comparisons between the approaches within both the mesh and point cloud methods, including mean differences, 95 CIs, and p-values. (*) indicates a statistically significant difference. N=17.*

Planned vs achieved planes Allograft		**Mesh method**	**Point cloud method**	**Mean difference (95% CI)**	**p-value**
Number of points		2480.7 ± 2114.6	26.1 ± 17.5	-	-
Distance Point-to-Point (in mm)	Mean ± std	-0.23 ± 1.60	0.47 ± 2.32	-0.70 [-1.54,0.14]	0.10
	Min	-3.52	-3.96		
	Max	5.07	6.72		
	Median	-0.44	0.36		
Distance COG-to-plane (in mm)	Mean ± std	-0.07 ± 0.26	-0.03 ± 0.15	-0.04 [-0.05,0.29]	0.15
	Min	-0.96	-0.29		
	Max	0.34	0.25		
	Median	-0.03	0.00		
Distance COG-to COG (in mm)	Mean ± std	4.22 ± 2.27	4.41 ± 2.44	-0.19 [-0.66,1.04]	0.64
	Min	0.20	0.84		
	Max	7.72	9.91		
	Median	4.17	3.75		
Pairwise comparisons					
Point-to-Point vs COG-to-plane		-0.08 [-0.93, 0.76], p=0.55	0.50 [-0.71,1.71], p=0.45		

COG-to-COG vs Point-to-Point	4.45[3.10,5.80], p<0.001 (*)	3.94 [2.51,5.37], p<0.001 (*)
COG-to-COG vs COG-to-plane	4.37 [3.09,5.64], p<0.001 (*)	4.44 [3.11,5.77], p<0.001 (*)

Table B3: Linear results using point-to-point distances of the mesh and point cloud methods in patient cases. N=28.

Patient #	Amount of cutting planes	Method	Mean ± std (mm)	Min	Max	Median
1	4	Mesh	-0.81 ± 0.41	-1.48	-0.43	-0.66
		Point cloud	-1.35 ± 0.27	-1.18	-1.08	-1.26
2	2	Mesh	-3.89 ± 0.72	-4.61	-3.17	-3.89
		Point cloud	-3.14 ± 0.88	-4.02	-2.25	-3.14
3	4	Mesh	0.99 ± 0.74	-0.02	2.05	0.96
		Point cloud	-0.45 ± 1.18	-1.40	1.58	-0.98
4	3	Mesh	0.26 ± 0.23	0.07	0.58	0.12
		Point cloud	-0.88 ± 1.01	-1.80	0.53	-1.36
5	5	Mesh	0.24 ± 0.85	-0.89	1.32	0.69
		Point cloud	-0.05 ± 0.99	-1.49	1.58	-0.16
6	1	Mesh	1.64	1.64	1.64	1.64
		Point cloud	0.59	0.59	0.59	0.59
7	3	Mesh	-0.64 ± 3.45	-3.22	4.24	-2.94
		Point cloud	0.01 ± 4.19	-3.10	5.92	-2.81
8	1	Mesh	-1.45	-1.45	-1.45	-1.45
		Point cloud	-1.55	-1.55	-1.55	-1.55
9	3	Mesh	-3.76 ± 0.57	-4.25	-2.95	-4.07
		Point cloud	-2.72 ± 0.29	-3.11	-2.43	-2.60
10	2	Mesh	1.87± 0.62	1.26	2.49	1.87
		Point cloud	2.00 ± 0.15	1.85	2.16	2.00

Table B4: Linear results using point-to-point distances of the mesh and point cloud methods in allograft cases. N=17.

Patient #	Amount of cutting planes	Method	Mean ± std	Min	Max	Median
1	4	Mesh	-0.48 ± 0.56	-1.15	0.07	-0.43
		Point cloud	0.14 ± 0.63	-0.91	0.71	0.38
2	1	Mesh	-0.55	-0.55	-0.55	-0.55
		Point cloud	0.17 ±	0.17	0.17	0.17
3	4	Mesh	-0.12 ± 0.65	-0.75	0.88	-0.31
		Point cloud	0.91±0.65	-0.19	1.45	1.18
5	5	Mesh	0.13 ± 2.76	-3.12	4.72	-0.29
		Point cloud	0.78±2.57	-1.83	5.26	0.32
7	3	Mesh	-0.53 ± 0.65	-1.35	0.23	-0.46
		Point cloud	-0.09±4.19	-3.20	5.82	-2.90

C – Angular outcomes

Table C1: Comparison of the mesh and point cloud methods across two quantification approaches, and resulting angular outcome measures in patient cases. N=28.

Planned vs achieved planes Patient		**Mesh method**	**Point cloud method**	**Mean difference (95% CI)**	**p-value**
Number of points		2771.9 ± 1964.5	27.8 ± 25.8	-	-
Angle between normal vectors (in degrees)	Mean ± std	6.60 ± 6.50	6.13 ± 5.43	0.47 [-1.07,2.01]	0.53
	Min	1.39	0.46		
	Max	28.44	26.74		
	Median	4.64	4.36		
Angle in roll (in degrees)	Mean ± std	-0.14 ± 4.01	-0.21 ± 3.88	-0.07 [-1.12,0.98]	0.89
	Min	-6.79	-6.88		
	Max	15.90	13.77		
	Median	-0.13	-0.25		
Angle in pitch (in degrees)	Mean ± std	-0.63 ± 6.84	0.15 ± 5.36	0.79 [-1.64,3.22]	0.51
	Min	-23.62	-19.05		
	Max	6.19	8.90		
	Median	1.99	1.54		

Table C2: Comparison of the mesh and point cloud methods across two quantification approaches, and resulting angular outcome measures in allograft cases. N=17.

Planned vs achieved planes Allograft		**Mesh method**	**Point cloud method**	**Mean difference (95% CI)**	**p-value**
Number of points		2480 ± 2114.6	26.1 ± 17.5		-
Angle between normal vectors (in degrees)	Mean ± std	5.33 ± 5.07	4.71 ± 2.92	0.62 [-1.78,3.02]	0.59
	Min	1.21	0.46		
	Max	18.22	9.87		
	Median	3.38	3.64		
Angle in roll (in degrees)	Mean ± std	0.03 ± 1.60	0.70 ± 2.30	-0.67 [-1.67,0.33]	0.17
	Min	-2.95	-2.80		
	Max	4.02	7.57		
	Median	-0.19	0.43		
Angle in pitch (in degrees)	Mean ± std	0.94 ± 7.08	2.91 ± 3.85	-1.96 [-4.63,0.70]	0.14
	Min	-17.79	-3.88		
	Max	18.07	9.36		
	Median	1.16	3.21		

Table C3: Angular results using point-to-point distances of the mesh and point cloud methods in patient cases. N=28.

Patient #	Amount of cutting planes	Method	Mean ± std (mm)	Min	Max	Median
1	4	Mesh	3.39 ± 1.00	2.33	4.89	3.17
		Point cloud	3.23 ± 0.92	2.03	4.60	3.14
2	2	Mesh	4.71 ± 0.25	4.46	4.96	4.71
		Point cloud	4.99 ± 0.75	4.24	5.75	4.99
3	4	Mesh	7.55 ± 1.80	5.93	10.51	6.88
		Point cloud	5.01 ± 2.62	2.52	9.02	4.26
4	3	Mesh	1.81 ± 0.50	1.39	2.52	1.51

		Point cloud	8.02 ± 2.73	4.18	10.23	9.66
5	5	Mesh	5.29 ± 2.01	2.86	8.26	4.99
		Point cloud	4.75 ± 1.64	2.14	6.99	4.41
6	1	Mesh	2.79	2.79	2.79	2.79
		Point cloud	2.18	2.18	2.18	2.18
7	3	Mesh	3.43 ± 1.00	2.50	4.82	2.97
		Point cloud	3.18 ± 2.18	0.46	5.76	3.61
8	1	Mesh	3.22	3.22	3.22	3.22
		Point cloud	2.04	2.04	2.04	2.04
9	3	Mesh	23.95 ± 3.74	19.29	28.44	23.82
		Point cloud	17.84 ± 7.95	7.44	26.74	19.33
10	2	Mesh	5.94 ± 3.22	2.71	9.16	5.94
		Point cloud	6.60 ± 3.49	3.11	10.09	6.60

Table C4: Angular results using point-to-point distances of the mesh and point cloud methods in allograft cases. N=17.

Patient #	Amount of cutting planes	Method	Mean ± std	Min	Max	Median
1	4	Mesh	4.66 ± 1.35	3.38	6.87	4.20
		Point cloud	3.49 ± 1.73	1.35	6.18	3.21
2	1	Mesh	8.88	8.88	8.88	8.88
		Point cloud	9.87	9.87	9.87	9.87
3	4	Mesh	10.94 ± 7.27	3.32	18.22	11.11
		Point cloud	6.98 ± 2.24	3.64	9.32	7.48
5	5	Mesh	2.32 ± 0.66	1.21	3.08	2.48
		Point cloud	3.72 ± 2.62	1.87	8.78	2.37
7	3	Mesh	2.60 ± 1.62	1.27	4.88	1.64
		Point cloud	3.28 ± 2.18	0.46	5.76	3.61

D – Quality assessment

All cutting planes with a smallest eigenvalue above 10 mm were selected. Twenty-four cutting planes were selected in patient cases and 12 in allograft cases using the mesh method, based on quality assessment of angular outcomes. Each patient is visualized in a different color.

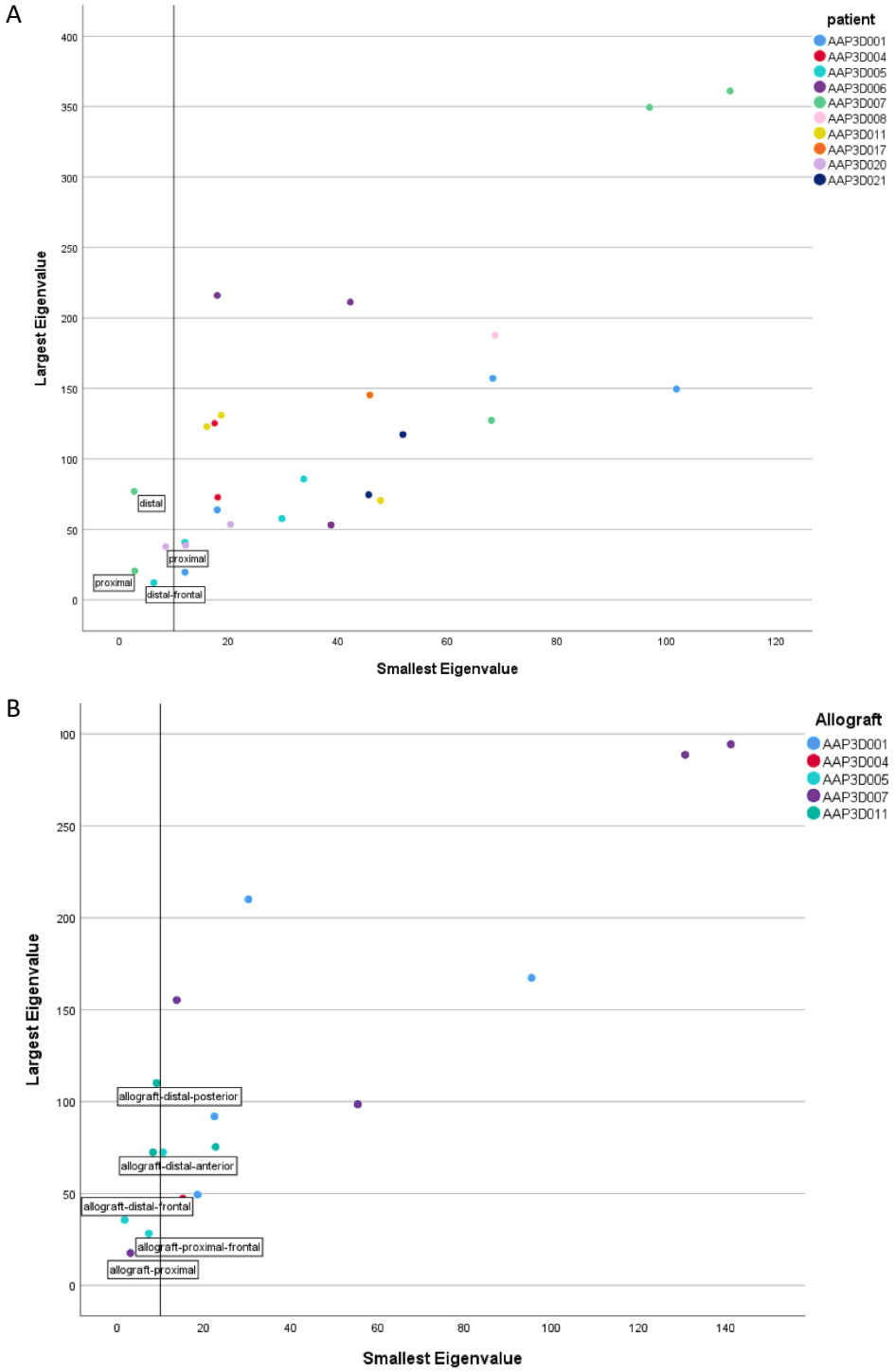
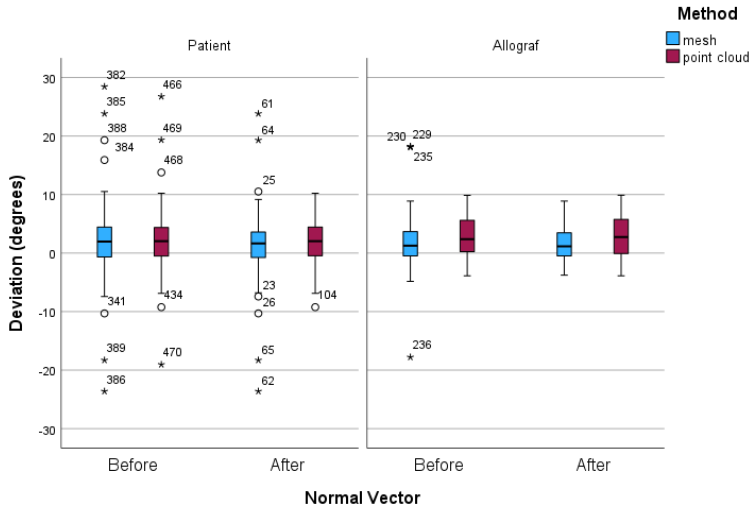
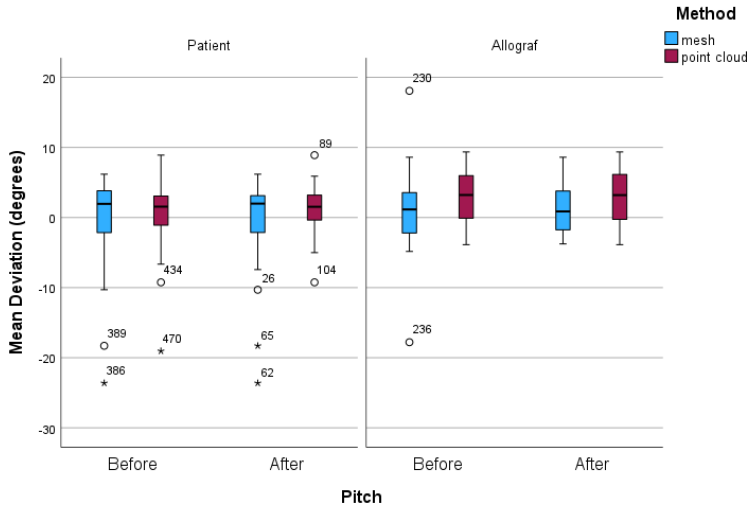


Figure D1: Scatter plots visualizing excluded cutting planes. Four planes were excluded in patient cases (A) and five planes in allograft cases (B). Each patient is visualized in a different color





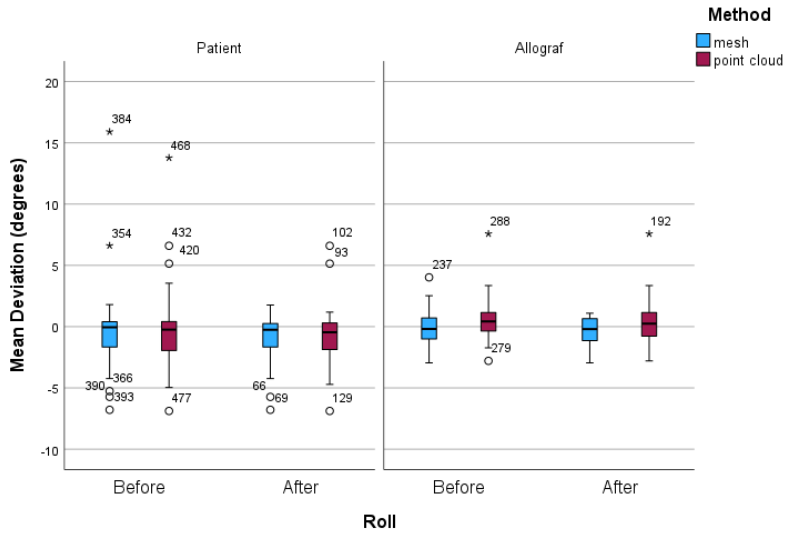


*Figure D2: Boxplots illustrating the mean angular deviation before and after quality assessment for normal vector approach, and pitch and roll approach. (*) indicates extreme outliers (>3*IQR. (o) indicates mild outliers (>1.5*IQR). The mesh method is presented in blue and the point cloud method in red.*

E – Flatness

Table E1: Comparison of the mesh and point cloud methods across two quantification approaches, and resulting flatness outcome measure in patient cases and resulting flatness outcome measure. (*) indicates a statistically significant difference. N=28.

Achieved plane Patient		**Mesh method**	**Point cloud method**	**Mean difference (95% CI)**	**p-value**
Corrected ISO-1101 standard	Mean ± std	2.78 ± 1.19	2.48 ± 2.14	0.30 [-0.54,1.15]	0.47
	Min	1.01	0.47		
	Max	5.60	9.66		
	Median	2.41	1.59		
Non corrected ISO-1101 standard	Mean ± std	4.33 ± 1.74	2.63 ± 2.30	1.70 [0.71,2.68]	0.002 (*)
	Min	1.48	0.44		
	Max	8.00	9.77		
	Median	3.94	1.62		
Pairwise comparisons					
Corrected vs non-corrected ISO-1101 standard		-1.55 [-1.86,-1.23] p < 0.001 (*)	-0.15 [-0.38,0.07] p = 0.17		

Table E2: Comparison of the mesh and point cloud methods across two quantification approaches, and resulting flatness outcome measure in allograft cases and resulting flatness outcome measure. (*) indicates a statistically significant difference. N=17.

Achieved plane Allograft		**Mesh method**	**Point cloud method**	**Mean difference (95% CI)**	**p-value**
Corrected ISO-1101 standard	Mean ± std	2.36 ± 0.90	1.46 ± 0.62	0.90 [0.35, 1.46]	0.003 (*)
	Min	0.54	0.51		
	Max	3.83	2.74		
	Median	2.22	1.36		
Non corrected ISO-1101 standard	Mean ± std	3.29 ± 1.17	1.60 ± 0.69	1.67 [0.99, 2.35]	<0.001 (*)
	Min	1.48	0.31		
	Max	6.32	2.63		
	Median	3.24	1.34		
Pairwise comparisons					
Corrected vs non-corrected ISO-1101 standard		-0.90 [-1.24,-0.56] p < 0.001 (*)	-0.14 [-0.28, 0.01] p = 0.07		

F – Surgical Margin

Table F1: Comparison of the mesh and point cloud methods and resulting surgical margin. (*) indicates a statistically significant difference. N=23.

Planned vs achieved SM		**Mean ± std**	**Min**	**Max**	**Median**
Mesh method	Planned	10.0 ± 5.8	0.8	22.0	7.2
	Achieved	10.2 ± 5.8	1.5	23.4	8.0
	Margin diff	-0.2 ± 2.2	-5.4	4.3	0.0
Point cloud method	Achieved	11.1 ± 5.7	4.3	23.4	8.7
	Margin diff	-1.1 ± 2.2	-4.6	5.8	-1.2
Mean difference (95% CI)		0.9 [0.24, 1.56]			
P-value		0.01 (*)			

G – COG-COG vs COG-Plane

COG-COG approach results in a linear deviation of 10.23 mm, while the COG-to-plane approach results in a deviation of only -0.16 mm

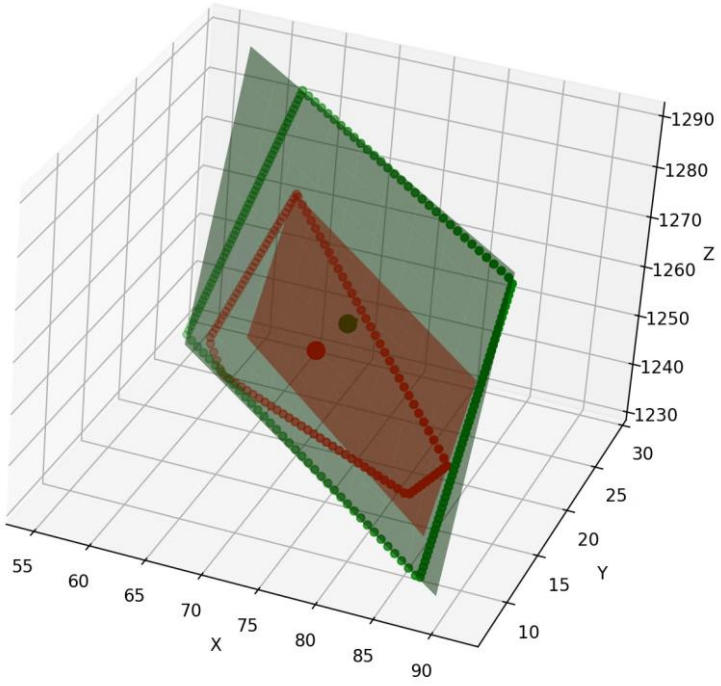
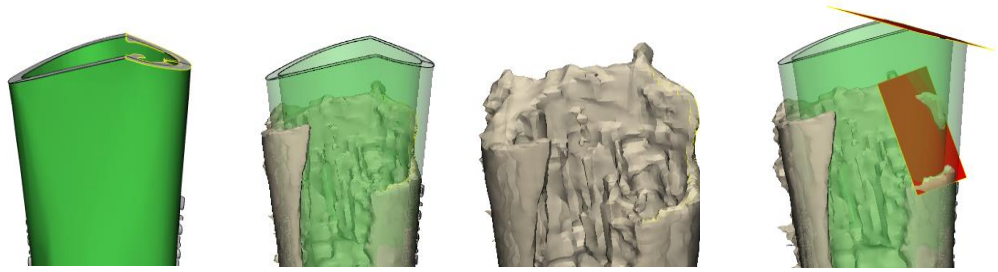


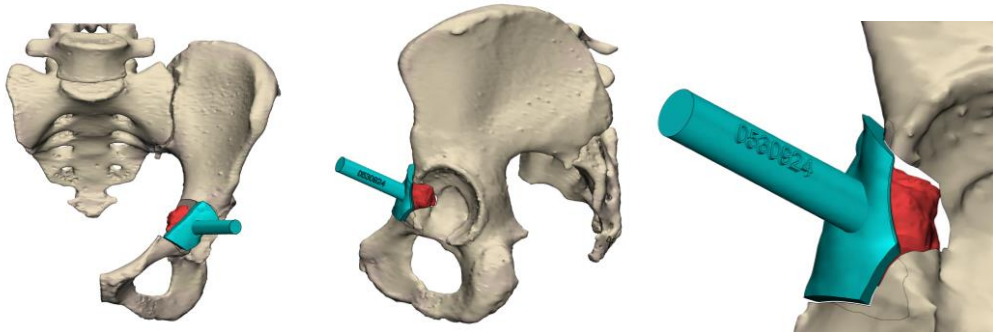
Figure G1: Planned plane and bone contour visualized in green. Achieved plane and bone contour visualized in red.

H - Outliers

Example of excluded cutting planes in allograft case:



Example of excluded cutting plane in patient case:



User Guide

When the STLs (mesh method) and XML (point cloud method) files are created, the Python script can be run. This scripts calculates the linear (Point-to-Point, COG-to-COG, and COG-to-plane) and angular (normal vector, and pitch and roll) outcome measures, flatness of the achieved plane (corrected and non-corrected ISO-standard). The surgical margin is calculated using another script. The results are exported to an Excel sheet.

1. Start by opening *XML_to_CSV.py files* (in MSc\Python code)
    a. Add filepath of new XML files in **filepaths** and change output_dir to the location where the CSV file should be saved.
    b. Run the script
2. Now, open *functions.py*
    a. Add the filepaths of the new STL and CSV files in the corresponding **allografts** or **patients** directory
    b. Run the script
    c. The results will be exported to an Excel file called results_patients.csv or results_allografts.csv. The Excel files will be saved in MSc\Results
    d. The Excel file has the following columns:
        i. Patient: contains patient code
        ii. Plane: contains specific cutting plane
        iii. Method: contains type of method (mesh/pointcloud)
        iv. Mean_Distance: contains the mean distance between points of achieved versus planned plane (mm)
        v. N_points_plane_fitting: contains number of points used for plane fitting
        vi. COG_COG_Distance: contains the distance between COG_planned and COG_achieved (mm)
        vii. COG_Plane_Distance: contains the average distance between COG_planned and achieved plane, and vise versa (mm)
        viii. Angle-NV: contains the angle between planned and actual normal vectors (degrees)
        ix. Roll: contains the angle between planned and actual normal vectors in medial-lateral direction (degrees)
        x. Pitch: contains the angle between planned and actual normal vectors in cranial-caudal direction (degrees)
        xi. Pitch-After: contains the pitch angle after quality assessment (degrees)
        xii. Roll-After: contains the roll angle after quality assessment (degrees)
        xiii. NV-After: contains the angle between normal vectord after quality assessment (degrees)
        xiv. Flatness_ISO: contains the uncorrected flatness ISO-1101 standard (mm)
        xv. Distance_between_planes: contains the corrected flatness ISO-1101 standard (mm)
3. Next, open *surgical_margin.py*
    a. Add the filepaths of the new STL and CSV files in the **patients** directory
    b. Run the script

c. The results will be exported to an Excel file called results_surgical_margin.csv. The Excel file will be saved in MSc\Results
d. The Excel file has the following columns:
    i. Patient: contains patient code
    ii. Plane: contains specific cutting plane
    iii. Method: contains type of method (mesh/pointcloud)
    iv. Planned SM: contains the planned margin
    v. Actual SM: contains the achieved surgical margin
    vi. Margin difference: contains the surgical margin difference between the planned and achieved cutting plane

References

1. Hak DJ, Rose J, Stahel PF. Preoperative planning in orthopedic trauma: benefits and contemporary uses. . Orthopedics. 2010.
2. Mussi E, Mussa F, Santarelli C, Scagnet M, Uccheddu F, Furferi R, et al. Current Practice in Preoperative Virtual and Physical Simulation in Neurosurgery. Bioengineering 2020.
3. Sodhi N, Anis HK, Coste M, Ehiorobo JO, Chee A, Freund B, et al. A Nationwide Analysis of Preoperative Planning on Operative Times and Postoperative Complications in Total Knee Arthroplasty. J Knee Surg. 2019;32(11):1040–5.
4. Essert C, Joskowicz L. Chapter 32 - Image-based surgery planning. In: Zhou SK, Rueckert D, Fichtinger G, editors. Handbook of Medical Image Computing and Computer Assisted Intervention: Academic Press; 2020. p. 795–815.
5. Portnoy Y, Koren J, Khoury A, Factor S, Dadia S, Ran Y, et al. Three-dimensional technologies in presurgical planning of bone surgeries: current evidence and future perspectives. International journal of surgery. 2023.
6. Evrard R, Schubert T, Paul L, Docquier PL. Quality of resection margin with patient specific instrument for bone tumor resection. J Bone Oncol. 2022;34:100434.
7. Hammoudeh JA, Howell LK, Boutros S, Scott MA, Urata MM. Current Status of Surgical Planning for Orthognathic Surgery: Traditional Methods versus 3D Surgical Planning. Plastic and reconstructive surgery Gloabl open. 2015.
8. Wade RH, Kevu J, J. D. Pre-operative planning in orthopaedics: a study of surgeons' opinions. Injury. 1998.
9. Cartiaux O, Paul L, Docquier P-L, Raucent B, Dombre E, Banse X. Computer-Assisted and Robot-Assisted Technologies to Improve Bone-Cutting Accuracy When Integrated with a Freehand Process Using an Oscillating Saw. JBJS. 2010;92(11):2076–82.
10. Jeys L, Matharu GS, Nandra RS, Grimer RJ. Can computer navigation-assisted surgery reduce the risk of an intralesional margin and reduce the rate of local recurrence in patients with a tumour of the pelvis or sacrum? The Bone & Joint Journal. 2013;95-B(10):1417–24.
11. Khan F, Pearle A, Lightcap C, Boland PJ, Healey JH. Haptic robot-assisted surgery improves accuracy of wide resection of bone tumors: a pilot study. Clin Orthop Relat Res. 2013;471(3):851–9.
12. Wong K-C, Niu X, Xu H, Li Y, Kumta S. Computer Navigation in Orthopaedic Tumour Surgery. In: Zheng G, Tian W, Zhuang X, editors. Intelligent Orthopaedics: Artificial Intelligence and Smart Image-guided Technology for Orthopaedics. Singapore: Springer Singapore; 2018. p. 315–26.
13. Wong K-C, Sze K-Y, Wong IO-L, Wong C-M, Kumta S-M. Patient-specific instrument can achieve same accuracy with less resection time than navigation assistance in periacetabular pelvic tumor surgery: a cadaveric study. International Journal of Computer Assisted Radiology and Surgery. 2016;11(2):307–16.
14. Sallent A, Vicente M, Reverté MM, Lopez A, Rodríguez-Baeza A, Pérez-Domínguez M, et al. How 3D patient-specific instruments improve accuracy of pelvic bone tumour resection in a cadaveric study. Bone & joint research. 2017.
15. Benignus C, Buschner P, Meier M, Wilken F, Rieger J, Beckmann J. Patient Specific Instruments and Patient Individual Implants-A Narrative Review. Journal of personalized medicine. 2023.
16. Han S-B, Kim H-J, Kim T-K, In Y, Oh K-J, Koh I-J, et al. Computer navigation is effective in reducing blood loss but has no effect on transfusion requirement following primary total knee arthroplasty: a meta-analysis. Knee Surgery, Sports Traumatology, Arthroscopy. 2016;24(11):3474–81.
17. T.Ozaki, Flege S, Kevric M, Lindner N, Maas R, Delling G, et al. Osteosarcoma of the pelvis: experience of the Cooperative Osteosarcoma Study Group. Journal of Clinical Oncology. 2003.
18. Sluga M, Windhager R, Lang S, Heinzl H, Krepler P, Mittermayer F, et al. The Role of Surgery and Resection Margins in the Treatment of Ewing's Sarcoma. Clinical Orthopaedics and Related Research®. 2001;392:394–9.

19. Müller DA, Stutz Y, Vlachopoulos L, Farshad M, Fürnstahl P. The Accuracy of Three-Dimensional Planned Bone Tumor Resection Using Patient-Specific Instrument. Cancer Manag Res. 2020;12:6533–40.
20. Hoffmann C, Ahrens S, Dunst J, Hillmann A, Winkelmann W, Craft A, et al. Pelvic Ewing sarcoma: a retrospective analysis of 241 cases. Cancer 1999.
21. Han I, Lee YM, Cho HS, Oh JH, Lee SH, Kim HS. Outcome after Surgical Treatment of Pelvic Sarcomas. Clinics in Orthopedic Surgery. 2010.
22. Gradinger R, Rechl H, Hipp E. Precision tumour resection and reconstruction using image-guided computer navigation. Clinical Orthopaedics and Related Research. 2007.
23. Gouin F, Paul L, Odri GA, Cartiaux O. Computer-Assisted Planning and Patient-Specific Instruments for Bone Tumor Resection within the Pelvis: A Series of 11 Patients. Sarcoma. 2014;2014:842709.
24. Fuchs B, Hoekzema N, Larson DR, Inwards CY, Sim FH. Osteosarcoma of the Pelvis: Outcome Analysis of Surgical Treatment. Clinical Orthopaedics and Related Research®. 2009;467(2):510–8.
25. Avedian RS, Haydon RC, Peabody TD. Multiplanar Osteotomy with Limited Wide Margins: A Tissue Preserving Surgical Technique for High-grade Bone Sarcomas. Clinical Orthopaedics and Related Research®. 2010;468(10).
26. Sternheim A, Daly M, Qiu J, Weersink R, Chan H, Jaffray D, et al. Navigated pelvic osteotomy and tumor resection: a study assessing the accuracy and reproducibility of resection planes in Sawbones and cadavers. J Bone Joint Surg Am. 2015;97(1):40–6.
27. Ritacco LE, Milano FE, Farfalli GL, Ayerza MA, Muscolo DL, Aponte-Tinao LA. Accuracy of 3-D planning and navigation in bone tumor resection. Orthopedics. 2013;36(7):e942–50.
28. Ackermann J, Liebmann F, Hoch A, Snedeker JG, Farshad M, Rahm S, et al. Augmented Reality Based Surgical Navigation of Complex Pelvic Osteotomies-A Feasibility Study on Cadavers. Appl Sci-Basel. 2021;11(3):19.
29. Bai G, He D, Yang C, Lu C, Huang D, Chen M, et al. Effect of digital template in the assistant of a giant condylar osteochondroma resection. J Craniofac Surg. 2014;25(3):e301–4.
30. Brouwer de Koning SG, Ter Braak TP, Geldof F, van Veen RLP, van Alphen MJA, Karssemakers LHE, et al. Evaluating the accuracy of resection planes in mandibular surgery using a preoperative, intraoperative, and postoperative approach. Int J Oral Maxillofac Surg. 2021;50(3):287–93.
31. Caiti G, Dobbe JGG, Strackee SD, van Doesburg MHM, Strijkers GJ, Streekstra GJ. A 3D printed cast for minimally invasive transfer of distal radius osteotomy: a cadaver study. Int J Comput Assist Radiol Surg. 2021;16(3):505–13.
32. Cartiaux O, Paul L, Francq BG, Banse X, Docquier PL. Improved accuracy with 3D planning and patient-specific instruments during simulated pelvic bone tumor surgery. Ann Biomed Eng. 2014;42(1):205–13.
33. Chan HHL, Sahovaler A, Daly MJ, Ferrari M, Franz L, Gualtieri T, et al. Projected cutting guides using an augmented reality system to improve surgical margins in maxillectomies: A preclinical study. Oral Oncology. 2022;127:7.
34. Cho HS, Park MS, Gupta S, Han I, Kim HS, Choi H, et al. Can Augmented Reality Be Helpful in Pelvic Bone Cancer Surgery? An In Vitro Study. Clin Orthop Relat Res. 2018;476(9):1719–25.
35. De Stavola L, Cristoforetti A, Fincato A, Nollo G, Ghensi P, Cantarutti A, et al. Accuracy and Technical Predictability of Computer Guided Bone Harvesting from the Mandible: A Cone-Beam CT Analysis in 22 Consecutive Patients. J Funct Biomater. 2022;13(4).
36. Dobbe JG, Kievit AJ, Schafroth MU, Blankevoort L, Streekstra GJ. Evaluation of a CT-based technique to measure the transfer accuracy of a virtually planned osteotomy. Med Eng Phys. 2014;36(8):1081–7.
37. Fujiwara T, Kunisada T, Takeda K, Hasei J, Nakata E, Nakahara R, et al. Intraoperative O-arm-navigated resection in musculoskeletal tumors. J Orthop Sci. 2018;23(6):1045–50.
38. Gerbers JG, Ooijen PM, Jutte PC. Computer-assisted surgery for allograft shaping in hemicortical resection: a technical note involving 4 cases. Acta Orthopaedica. 2013.

39. Hasan W, Daly MJ, Chan HHL, Qiu J, Irish JC. Intraoperative cone-beam CT-guided osteotomy navigation in mandible and maxilla surgery. Laryngoscope. 2020;130(5):1166–72.
40. Hoving AM, Kraeima J, Schepers RH, Dijkstra H, Potze JH, Dorgelo B, et al. Optimisation of three-dimensional lower jaw resection margin planning using a novel Black Bone magnetic resonance imaging protocol. PLoS One. 2018;13(4):e0196059.
41. Liu Z, Zhong Y, Lyu X, Zhang J, Huang M, Liu S, et al. Accuracy of the modified tooth-supported 3D printing surgical guides based on CT, CBCT, and intraoral scanning in maxillofacial region: A comparison study. J Stomatol Oral Maxillofac Surg. 2024;125(5s2):101853.
42. Modabber A, Ayoub N, Redick T, Gesenhues J, Kniha K, Möhlhenrich SC, et al. Comparison of augmented reality and cutting guide technology in assisted harvesting of iliac crest grafts - A cadaver study. Ann Anat. 2022;239:151834.
43. Otsuki B, Takemoto M, Kawanabe K, Awa Y, Akiyama H, Fujibayashi S, et al. Developing a novel custom cutting guide for curved peri-acetabular osteotomy. Int Orthop. 2013;37(6):1033–8.
44. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Jaworowski J, et al. Supporting fibula free flap harvest with augmented reality: A proof-of-concept study. Laryngoscope. 2020;130(5):1173–9.
45. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Mazurek M, et al. Supporting mandibular resection with intraoperative navigation utilizing augmented reality technology - A proof of concept study. J Craniomaxillofac Surg. 2019;47(6):854–9.
46. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Mazurek M, et al. Navigation-guided fibula free flap for mandibular reconstruction: A proof of concept study. J Plast Reconstr Aesthet Surg. 2019;72(4):572–80.
47. Pietruski P, Majak M, Świątek-Najwer E, Żuk M, Popek M, Świecka M, et al. Replacing cutting guides with an augmented reality-based navigation system: A feasibility study in the maxillofacial region. Int J Med Robot. 2023;19(3):e2499.
48. Qu M, Hou Y, Xu Y, Shen C, Zhu M, Xie L, et al. Precise positioning of an intraoral distractor using augmented reality in patients with hemifacial microsomia. J Craniomaxillofac Surg. 2015;43(1):106–12.
49. Siegel MA, Balach T, Sweeney KR, Nystrom LM, Colman MW. Sacroiliac joint cut accuracy: Comparing new technologies in an idealized sawbones model. J Surg Oncol. 2020;122(6):1218–25.
50. Sun M, Lin L, Chen X, Xu C, Zin MA, Han W, et al. Robot-assisted mandibular angle osteotomy using electromagnetic navigation. Ann Transl Med. 2021;9(7):567.
51. Ter Braak TP, Brouwer de Koning SG, van Alphen MJA, van der Heijden F, Schreuder WH, van Veen RLP, et al. A surgical navigated cutting guide for mandibular osteotomies: accuracy and reproducibility of an image-guided mandibular osteotomy. Int J Comput Assist Radiol Surg. 2020;15(10):1719–25.
52. Vanesa V, Irene MP, Marta AS, Francisco José PF, Miguel BS, Mireia RM, et al. Accuracy of virtually planned mandibular distraction in a pediatric case series. J Craniomaxillofac Surg. 2021;49(2):154–65.
53. Weijs WL, Coppen C, Schreurs R, Vreeken RD, Verhulst AC, Merkx MA, et al. Accuracy of virtually 3D planned resection templates in mandibular reconstruction. J Craniomaxillofac Surg. 2016;44(11):1828–32.
54. Winnand P, Ayoub N, Redick T, Gesenhues J, Heitzer M, Peters F, et al. Navigation of iliac crest graft harvest using markerless augmented reality and cutting guide technology: A pilot study. Int J Med Robot. 2022;18(1):e2318.
55. Wong KC, Kumta SM. Joint-preserving tumor resection and reconstruction using image-guided computer navigation. Clin Orthop Relat Res. 2013;471(3):762–73.
56. Wong KC, Kumta SM, Geel NV, Demol J. One-step reconstruction with a 3D-printed, biomechanically evaluated custom implant after complex pelvic tumor resection. Comput Aided Surg. 2015;20(1):14–23.

57. Zhao Z, Zhang Y, Lin L, Huang W, Xiao C, Liu J, et al. Intelligent electromagnetic navigation system for robot-assisted intraoral osteotomy in mandibular tumor resection: a model experiment. Front Immunol. 2024;15:1436276.
58. van der Kooij M, Gerbers JG, Kaptein BL. Outcome measures to quantify the differences between planned and actual osteotomy planes in computer-assisted surgery: a scoping review. 2024.
59. Standardization. IOf. ISO 1101:2004: Geometrical product specifications (GPS) — Geometrical tolerancing — Tolerances of form, orientation, location and run-out. Geneva, Switzerland2004.