PRE-OPERATIVE PLANNING FOR SHOULDER ARTHROPLASTY
AN EVALUATION OF DIFFERENT PLANNING SOFTWARE SYSTEMS

Authors:
Tom Fournveau (01504970)
Elias Van Haute (01503938)

Supervisors:
Prof. Dr. Lieven De Wilde
Prof. Dr. Alexander Van Tongel
Dr. Hans Van der Bracht

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Medicine in Medicine

Academic year: 2018 - 2020
This page is not available because it contains personal information.
Ghent University, Library, 2021.
Preface

Dear reader,

This dissertation is the result of months of work in the interesting domain that is orthopaedics. When I started working on this subject, I was blissfully unaware of the many aspects that are involved when planning an orthopaedic surgery. The first few months were spent learning to work with new software, which was a challenge, but very rewarding once I got the hang of it. Soon, Elias and I were able to come up with our own measuring techniques while being guided in the right direction by Prof. Dr. De Wilde and Prof. Dr. Van Tongel. They always took the time to carefully explain difficult topics, even when they were on a tight schedule, for which I am very grateful. Their enthoulsim about this study was the nudge in the back that I needed to deliver a result that would make them proud.

The second person I would like to thank is Dr. Hans Van der Bracht for helping us collect the data that we needed to complete this study. The prospect of contacting big medical companies was quite daunting as a young and inexperienced medical student. Dr. Van der Bracht taught me valuable insights in what to do and what not to do when contacting these large companies. He also laid the foundation for this study by collecting info on how these medical companies performed their measurements, which was a big help once we gathered all our data.

To say this study has sparked my interest in orthopaedics is an understatement. I found the subject to be extremely interesting, full of little aspects that remain unexplored and/or unsolved. I plan on continuing the research on the subject.

I wish you a lot of reading pleasure and hope you learn something new!

Tom Fournear

Ghent, 11/12/2019
Abstract

Background: The correct placement of components in total shoulder arthroplasty and reverse shoulder arthroplasty is essential for a good outcome (optimal stability, optimal function, avoid complications, ...). In order to acquire optimal positioning of the components, previous studies have shown that preoperative planning in shoulder arthroplasty should be performed on 3D reconstructed glenoids. However, no consensus exists regarding the methods that can be used to perform these measurements. Over the past years, multiple medical companies have emerged who have each developed their own method of performing glenoid measurement.

Objectives: The purpose of this study was to compare multiple different methods used by different companies and the investigators and to determine whether or not they resulted in the same measurements for inclination and version of the glenoid.

Methods: CT-scans were made of 9 different cadavers, 17 shoulders were taken into account, 1 shoulder was excluded from this study. All CT-scans were sent to the companies who had agreed to participate in this study. Each company was expected to measure the glenoid inclination and version, using their self-developed methods. A survey was also sent to these companies, with questions regarding the method they used. Furthermore, measurements of these shoulders were performed by the researchers, using 2 different methods (IG-plane method and SM-plane method).

The measurements of inclination were then statistically analyzed, first by using a Friedman ANOVA test to determine statistically significant differences in measurements for inclination between all methods, followed by post-hoc Wilcoxon tests to determine differences in a pairwise manner. This was repeated for the version measurements.

Results: Statistical analysis of the measurements for inclination and version showed significant differences between almost all methods. However, this discrepancy was more prominent for inclination than for version. Out of 15 pairs, 10 pairs differed significantly for inclination (p<0.008). Only 4 pairs differed significantly for version (p<0.008).

Conclusion: The findings of this study did not support the hypothesis that the different methods would produce measurements that do not statistically differ. These differences were most prominent for inclination. This means however, that the components would be differently positioned in the same shoulder, depending on which company performed the pre-operative
measurements. This leads to believe that certain methods can be superior to others. Clinical relevance and reproducibility should be taken into account when deciding which method to use.

Samenvatting

**Achtergrond:** Een correcte positionering van de componenten bij TSA en RSA is essentieel voor een optimale uitkomst (optimale stabiliteit, optimale functie, vermijden van complicaties, …). Om optimale positionering van de componenten te bekomen, werd in vroegere studies aangetoond dat pre-operatieve planning moet gebeuren, gebruik makend van 3D reconstructies van het glenoid. Bij gebrek aan consensus omtrent de methodes die gebruikt moeten worden om deze metingen uit te voeren, zijn er verschillende medische bedrijven die elk hun eigen methode ontwikkeld hebben.

**Doel:** Het doel van deze studie bestond erin de methodes van de verschillende bedrijven met elkaar te vergelijken, met als doel het testen van de hypothese dat al deze methodes dezelfde resultaten bij meting zouden geven.

**Methodes:** Er werd gebruik gemaakt van 17 schouders van 9 verschillende kadavers. Elke schouder werd door elk bedrijf opgemeten, telkens gebruik makend van hun zelf ontwikkelde methode. Daarnaast werd aan elk bedrijf informatie gevraagd omtrent de methodes die ze gebruikten, om eventueel gevonden verschillen te verklaren. Bovendien werd door de onderzoekers elke schouder gemeten op twee verschillende manieren (IG-plane en SM-plane methodes). Voor statistische analyse werd gebruik gemaakt van een Friedman ANOVA test voor inclinatie en versie, gevolgd door post-hoc paargewijze vergelijking met Wilcoxon tests.

**Resultaten:** Na statistische analyse bleken er duidelijk statistisch significante verschillen te bestaan tussen de verschillende methodes. Deze waren vooral prominent voor inclinatie, maar ook opvallend aanwezig voor versie.

**Conclusie:** De hypothese dat de verschillende methodes geen statistisch significante verschillen zouden voortbrengen werd aldus verworpen. Rekening houdend met de belangrijke klinische gevolgen, kan er vanuit gegaan worden dat bepaalde methodes de voorkeur moeten genieten.
1. Introduction

TSA (total shoulder arthroplasty) and RSA (reverse shoulder arthroplasty) are both effective treatments for different types of shoulder pathology. They are useful in both pain relief as well as function repair.\(^1\) A correct and accurate placement of the different components in both TSA and RSA is essential in achieving a good outcome. A correct positioning of the components is essential to 1) obtain good primary stability which ensures faster rehabilitation, 2) allow for optimal function and 3) avoid complications in TSA and RSA. Studies show that in TSA, an excessive retroversion and superior inclination are associated with higher frequencies of clinical failure.\(^2\) Clinical failure is often a result of loosening of the components. This is determined by multiple risk factors, such as the implantation technique, glenoid component type, rotator cuff status and posterior subluxation of the humeral head. The latter can be explained by the eccentric force acting on the glenoid component which leads to the occurrence of “rocking horse” phenomenon.\(^{1-5}\) Also, in RSA, an optimal component positioning is essential. An inferior position with slight inferior tilt (0-10°) is advised. Positioning the component too high will increase the risk of notching, instability and the occurrence of glenoid lucency.\(^5\) Along with the insight and experience of the surgeon, they are the corner stone of the procedure, which leads to better outcome in both the short and long term.\(^{1, 4, 6}\)

Multiple techniques are available to perform these measurements. Until recent years, 2 dimensional (2D) computed tomography (CT) was the standard imaging on which these measurements were performed.\(^2, 4, 7\) The accuracy of this technique is strongly dependent on the relation between the scapular plane and the axis of the CT-scanner. This relation can vary greatly under the influence of the patient’s position and the in situ scapular orientation. This is alarming, given the fact that important preoperative decisions are made based on these measurements.\(^4\)

Newly emerging modern imagine techniques and 3D technology that are available at present, give us the ability to convert a 2D CT-image into a 3D model. This model can be freely moved and positioned within the available software. As a consequence, the patient’s position within the CT-scanner and the in situ scapular position are no longer of any relevance. Thus, the 3D model can be oriented as such that the measurements can be performed in an accurate and reproducible manner.\(^1, 3, 4, 8, 9\)
Most of the preoperative planning systems rely on the user to conduct the measurements manually on the 3D models. This process is often time-consuming and has a risk of user errors. In recent years, multiple medical companies that sell shoulder prosthesis have started to conduct this process, commissioned by the surgeon. This raises the question whether or not these measurements are sufficiently accurate and whether or not they are easily reproducible. Multiple medical companies employ a self-developed technique in measuring inclination and version.

The goal of this study is to make 1) a comparison between the measurements conducted by the different companies (using the same CT-scans), 2) to compare these measurements to the ones made manually by the investigators (also using the same CT-scans) and 3) try to explain possible differences based on different measurement techniques. Our hypothesis is that there will be no statistically significant differences between all the measurements (using the same CT-scans).

2. Materials and Methods

Bilateral CT-scans were acquired of 9 cadavers. No selection was performed based on the status of the shoulder joint. One shoulder was excluded based on the fact that a shoulder prosthesis was already present, bringing the total number of unique shoulder joints to 17. All 17 shoulders were considered as being completely independent from each other.

A. Three-dimensional reconstruction

A 3D reconstruction was performed for 17 unique shoulders in a commercially available program (Mimics; Materialise). The Mimics software constructs a 3D model of the imported CT-scans, with a preset range of Hounsfield Units (HU). This range had to be manually adjusted by the investigators to achieve accurate 3D models of the glenohumeral joint. After the construction of the 3D model, the wrap function in Mimics was used to smoothen out any holes or irregularities that still existed within the models. Then, all 3D models were exported to 3Matic (Materialise) to conduct the necessary measurements.

In order to determine multiple essential points, axes and planes; a variety of functions provided by the 3Matic software were utilised.
B. Measurements by investigators

Two methods were used by the investigators. The Inferior Glenoid circle method was chosen because it mainly uses the inferior part of the glenoid in the determination of the glenoid plane and doesn’t take the tuberculum superior into consideration. This allowed the investigators to compare methods that used different reference points to determine the glenoid plane. The Superior Max method was chosen, based on the fact that most of the companies also used this method or a method that was similar.

In addition, both methods have been proven to be reproducible with good intra- and interobserver reliabilities.(6)

B.1. Determination of points and planes on the glenoid surface

1. Inferior Glenoid (IG) plane: the "mark" function was used to mark the antero-inferior glenoid rim (if necessary extended to also include the posterior rim, in case of distinct erosion of the anterior rim). The "mark"-function is used to mark all of the surface triangles that construct the 3D-structure. Subsequently the software was used to determine the mathematically best fitting circle for the marked structure. This circle was called the "inferior glenoid circle". The midpoint of the inferior glenoid circle was marked and given the name "IGm". The “Inferior glenoid plane” was defined in the software as the plane going through the midpoint "IGm", parallel to the "inferior glenoid circle", effectively constructing the plane that contains both the midpoint “IGm" as well as the "inferior glenoid circle” as these two naturally lay in the same plane.(6)

2. Superior Max (SM) plane: the "mark" function was used to mark the entire glenoid surface. Subsequently the border of the marked surface was separated, and a best fitting circle was constructed on this border. This circle was called the “Superior Max circle” and its midpoint “SMm”. The “Superior Max plane” was defined in the software as the plane going through the midpoint “SMm”, parallel to the “superior max circle”, effectively constructing the plane that contains both the midpoint “SMm” as well as the “superior max circle” as these two naturally lay in the same plane.(6)
B.2. Determination of the scapular reference plane and Cartesian coordinate system

The orientation of both the IG-plane and SM-plane is measured by using the scapular plane as a reference. The scapular plane is determined by 3 points: the centre of the glenoid circle, the most inferior point of the scapula (Sinf) and the most medial point of the scapula (trigonum scapulae). As the exact location of the trigonum scapulae is sometimes hard to determine in 3D models, the intersection between the extension of the spina scapulae and the medial border of the scapula was used. The scapular plane is unique for every type of glenoid plane, as every type of glenoid plane has its unique centre (IGm and SMm).

A Cartesian coordinate system was constructed specific for each glenoid with its origin in the midpoint (IGm or SMm for the inferior glenoid and superior max planes respectively) of the glenoid plane. The Y-axis was defined as the intersection of the glenoid plane and the scapular plane. The X-axis is constructed parallel to the glenoid plane and perpendicular to the Y-axis. The Z-axis is perpendicular to both the X-axis and the Y-axis.$^{(6, 10)}$
Fig. II

A: determination of the 3 points constructing the scapular reference plan (Sinf, trigonum scap. and IGm/SMm).
B: construction of the cartesian coordinate system with the glenoid centre (IGm/SMm) as origin. (example shown in B. uses the SM-plane method).

**B.3. Determination of the Spina scapular axis**

The Spina scapular axis (SSA) was defined as the line between the trigonum scapulae and the centre of the glenoid circle. As the SSA also utilises the centre of the glenoid circles, the SSA is unique for every type of glenoid circle.\(^6\)

**B.4. Measurements**

Version and inclination measurements were obtained for all 17 unique shoulders, utilising both the IG- and the SM-plane. Version was defined as the angle between the Z-axis and the scapular plane. Anteversion was reported as a positive angle, while retroversion was reported as a negative angle.

Inclination was determined by calculating the complementary angle of the angle between the Y-axis and the SSA, e.g. if the angle between the Y-axis and the SSA measures 84°, the complementary angle is calculated by subtracting this measurement from 90° (in this case: 90° - 84° = 6° of inclination). Superior inclination was reported as a positive angle, while inferior inclination was reported as a negative angle.
C. Measurements by medical companies
Each company (Medacta, Materialise, Tornier-Wright, Zimmer-Biomet, Arthrex) was sent a questionnaire in which they were asked to specify the methods they used to perform the measurements. They were asked about the determination of the glenoid centre, the scapular plane, the transverse/axial plane (or SSA) and the reference points used on the glenoid plane. Arthrex was asked twice to partake in this study and they refused both times.

D. Statistical analysis
Version and inclination measurements were compared using the Friedman ANOVA test. Wilcoxon tests were applied post hoc to determine differences in a pairwise manner. Bonferroni corrections were applied to compensate for multiple comparisons.
For continuous, non-parametrically distributed outcome measures, intraobserver and interobserver reliability were assessed by intraclass correlation coefficients (ICC) based on 2-way mixed effect models, using an absolute agreement definition. Two independent investigators measured all 17 shoulders thrice, for both the IG-plane method and the SM-plane method. Statistical analysis was performed using the IBM SPSS 26 software.

3. Results
A. Measurements by investigators
A.1. Intraobserver and interobserver variability
Intra- and interobserver reliability were assessed by intraclass correlation coefficients. An ICC value <0.50 indicates a poor reliability, ICC values between 0.50 and 0.75 indicate moderate reliability, ICC values between 0.75 and 0.90 indicate good reliability and ICC values >0.90 indicate excellent reliability. These results are reported as “ICC value average measures (single measures).”

The intraobserver reliability was excellent for both inclination and version using the IG-plane method, with ICC ranging from 0.901 to 0.971 (0.753 to 0.918) for inclination and from 0.942 to 0.947 (0.844 to 0.857) for version. The interobserver reliability was good for inclination and excellent for version with ICC 0.871 (0.772) and 0.930 (0.869) respectively using the IG-plane method.
The intraobserver and interobserver reliabilities were excellent for both inclination and version using the SM-plane method, with ICC ranging from 0.952 to 0.996 (0.936 to 0.989) for inclination and from 0.981 to 0.991 (0.944 to 0.972) for version.

A.2. Descriptive statistics

Descriptive statistics are presented as mean (95% confidence interval) and are displayed in Table II and Table III.

- UGent-IG: The mean inclination was 16.64° (13.28°,20°) and the mean version was -0.58° (-5.62°,4.46°).
- UGent-SM: The mean inclination was 6.28° (2.18°,10.38°) and the mean version was -1.24° (-6.4°,3.92°).

B. Methods of medical companies

- Medacta (MyShoulder) use 4 points to determine the glenoid plane. These points are the most anterior, superior, posterior and inferior points on the glenoid. It is unclear which point is used to define the glenoid centre. Three points are used to identify the scapular plane: glenoid centre, the angulus inferior, and the trigonum scapulae. The transverse axis is defined as the axis between the glenoid centre and trigonum scapulae. This method is similar to the SM-method used by the investigators. However, the definition of the glenoid centre is different to the definition of the glenoid centre in the SM-method. This can have an influence on the position of the glenoid centre and the orientation of the transverse axis compared to the SM-method.

- Materialise: utilise an automated software to perform PSI planning. This software mathematically determines the best fitting glenoid plane based on all points included in the glenoid fossa. The glenoid centre corresponds to the geometrical centre of the glenoid and is automatically determined. The scapular plane is determined by using 3 scapular points: glenoid centre, angulus inferior and trigonum scapulae. The transverse axis is once again defined by using the glenoid centre and the trigonum scapulae. This method is similar to the SM-method used by the investigators. However, the definition of the glenoid centre is different to the definition of the glenoid centre in the SM-method. This can have an influence on the position of the glenoid centre and the orientation of the transverse axis compared to the SM-method.
- **Tornier-Wright (Blueprint)**: utilise an automated software (Blueprint) to perform PSI planning. The software mathematically determines the best fitting glenoid plane based on all points included in the glenoid fossa. The glenoid centre is automatically determined by the software. It is unclear how Blueprint determines the centre, however it may be assumed that it corresponds with the geometrical centre. The scapular plane is also mathematically determined. The software uses all points of the scapula to construct a best fit plane. The transverse axis is determined by using the glenoid centre and the trigonum scapulae. This method is similar to the SM-method used by the investigators. However, the definition of the glenoid centre is different to the definition of the glenoid centre in the SM-method. This can have an influence on the position of the glenoid centre and the orientation of the transverse axis compared to the SM-method. The construction of the scapular plane is also different, as Blueprint constructs a best fitting plane on the entire scapula, while the SM-method constructs the scapular plane using 3 reference points (glenoid centre, trigonum scapulae and angulus inferior).

- **Zimmer-Biomet (PSI Shoulder System)**: determine the glenoid plane by using 3 points: the anterior, superior and posterior points. The method used to determine the glenoid centre is unclear. It can be assumed the glenoid centre is determined by the intersection of the anteroposterior and the inferosuperior axes. To determine the scapular plane, Zimmer-Biomet also use 3 points: the glenoid centre, the angulus inferior and the trigonum scapulae.

- **Arthrex (VIP)**: Arthrex decided not to cooperate, thus they were not included in this study.

*Table I Methods used by different companies and the investigators*

<table>
<thead>
<tr>
<th>Company/Method</th>
<th>Glenoid Centre</th>
<th>Glenoid Reference Points</th>
<th>Scapular Plane</th>
<th>Transverse Axis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tornier-Wright (Blueprint)</td>
<td>Geometric? (automatic)</td>
<td>All points of the glenoid fossa</td>
<td>Best fit plane of scap. (automatic)</td>
<td>Glenoid centre - trigonum scap.</td>
</tr>
</tbody>
</table>
### C. Measurements of companies

Descriptive statistics are presented as mean (95% confidence interval) and are displayed in Table II and Table III.

- **Medacta**: The mean inclination was 7.15° (2.95°,11.35°) and the mean version was -3.54° (-8.65°,1.57°).
- **Materialise**: The mean inclination was 5.38° (2.82°,7.94°) and the mean version was -1.15° (-6.13°,3.83°).
- **Tornier-Wright (Blueprint)**: The mean inclination was 10.31° (5.51°,15.11°) and the mean version was -1° (-6.2°,4.2°).
- **Zimmer-Biomet**: The mean inclination was 0.38° (-3.69°,4.45°) and the mean version was 0.54° (-3.77°,4.85°).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthrex (VIP)</td>
<td>?</td>
<td>?</td>
<td>3 points on scapula: not clearly specified</td>
<td>?</td>
</tr>
</tbody>
</table>
Table II Descriptive statistics for inclination

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UGent-IG</strong></td>
<td>13</td>
<td>16,64°</td>
<td>3,36°</td>
<td>12°</td>
<td>22°</td>
</tr>
<tr>
<td><strong>UGent-SM</strong></td>
<td>13</td>
<td>6,28°</td>
<td>4,10°</td>
<td>-2°</td>
<td>12°</td>
</tr>
<tr>
<td><strong>Medacta (MyShoulder)</strong></td>
<td>13</td>
<td>7,15°</td>
<td>4,20°</td>
<td>-1°</td>
<td>15°</td>
</tr>
<tr>
<td><strong>Materialise</strong></td>
<td>13</td>
<td>5,38°</td>
<td>2,56°</td>
<td>-1°</td>
<td>11°</td>
</tr>
<tr>
<td><strong>Tornier-Wright (Blueprint)</strong></td>
<td>13</td>
<td>10,31°</td>
<td>4,80°</td>
<td>0°</td>
<td>19°</td>
</tr>
<tr>
<td><strong>Zimmer-Biomet (PSI Shoulder System)</strong></td>
<td>13</td>
<td>0,38°</td>
<td>4,07°</td>
<td>-10°</td>
<td>6°</td>
</tr>
</tbody>
</table>

Table III Descriptive statistics for version

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UGent-IG</strong></td>
<td>13</td>
<td>-0,58°</td>
<td>5,04°</td>
<td>-8°</td>
<td>9°</td>
</tr>
<tr>
<td><strong>UGent-SM</strong></td>
<td>13</td>
<td>-1,24°</td>
<td>5,16°</td>
<td>-9°</td>
<td>11°</td>
</tr>
<tr>
<td><strong>Medacta (MyShoulder)</strong></td>
<td>13</td>
<td>-3,54°</td>
<td>5,11°</td>
<td>-14°</td>
<td>7°</td>
</tr>
<tr>
<td><strong>Materialise</strong></td>
<td>13</td>
<td>-1,15°</td>
<td>4,98°</td>
<td>-11°</td>
<td>10°</td>
</tr>
<tr>
<td><strong>Tornier-Wright (Blueprint)</strong></td>
<td>13</td>
<td>-1,00°</td>
<td>5,20°</td>
<td>-10°</td>
<td>10°</td>
</tr>
<tr>
<td><strong>Zimmer-Biomet (PSI Shoulder System)</strong></td>
<td>13</td>
<td>0,54°</td>
<td>4,31°</td>
<td>-7°</td>
<td>10°</td>
</tr>
</tbody>
</table>

**D. Comparison between investigators and companies**

**D.1. Inclination**

A total of 13 shoulders were taken into account when performing the Friedman ANOVA test for inclination, as 4 shoulders had incomplete data for 1 or more measurements. The Friedman test showed significant differences (P<0.001) in inclination measurements between the 6 methods used. 15 pairwise Wilcoxon post-hoc tests were performed. Bonferroni correction lowered the threshold P-value to 0.0083. P-values for each pair are displayed in Table IV.
Of the 15 pairwise Wilcoxon post-hoc tests, 10 resulted in a statistically significant outcome. A statistically significant difference was found when comparing UGent-IG pairwise with UGent-SM (p<0.001), Medacta (p<0.001), Materialise (p<0.001), Blueprint (p=0.008) and Zimmer-Biomet (p<0.001). Furthermore, comparison of UGent-SM with Zimmer-Biomet resulted in a statistically significant result (p<0.001), whereas comparison with Medacta (p=0.136), Materialise (p=0.224) and Blueprint (p=0.022) resulted in p-values > 0.0083 (thus being statistically insignificant). Comparing Medacta with Materialise (p=0.012) and Blueprint (p=0.021) also resulted in p-values > 0.0083, while the comparison with Zimmer-Biomet showed a statistically significant result (p<0.001). The pairwise comparison of Materialise to Blueprint and Medacta also resulted in a statistically significant difference, with p = 0.002 and p<0.001 respectively. The final pairwise Wilcoxon test compared Zimmer-Biomet and Blueprint. this also resulted in a statistically significant result with p = 0.001.

*Table IV P-values obtained for inclination by performing post hoc Wilcoxon tests on 15 pairs of methods.*

<table>
<thead>
<tr>
<th>Method 1</th>
<th>Method 2</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGent-IG</td>
<td>UGent-SM</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Medacta (MyShoulder)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Materialise</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>UGent-SM</td>
<td>Medacta (MyShoulder)</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>Materialise</td>
<td>0.224</td>
</tr>
<tr>
<td></td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Medacta (MyShoulder)</td>
<td>Materialise</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Materialise</td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.001</td>
</tr>
</tbody>
</table>
D.2. Version

The Friedman ANOVA test for version also took 13 shoulders into account, as 4 shoulders had incomplete data for 1 or more measurements. The Friedman test showed significant differences (P=0.003) in version measurements between the 6 methods used. 15 pairwise Wilcoxon post-hoc tests were then performed to identify the pairs that presented statistically significant differences. Bonferroni correction was performed, which lowered the threshold P-value to 0.0083. P-values for each pair are displayed in Table V.

Of the 15 pairwise Wilcoxon post-hoc tests, 4 resulted in p-values < 0.0083. The comparisons between UGent-IG and Medacta (p=0.006), between UGent-SM and Zimmer-Biomet (p=0.004), between Medacta and Materialise (p=0.002) and between Medacta and Zimmer-Biomet (p=0.003) were the only 4 comparisons with a statistically significant outcome. All other comparisons resulted in p-values >0.0083.

Table V P-values obtained for version by performing post hoc Wilcoxon tests on 15 pairs of methods.

<table>
<thead>
<tr>
<th>Method 1</th>
<th>Method 2</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGent-IG</td>
<td>UGent- SM</td>
<td>0.124</td>
</tr>
<tr>
<td></td>
<td>Medacta (MyShoulder)</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Materialise</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.550</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>0.379</td>
</tr>
<tr>
<td>UGent-SM</td>
<td>Medacta (MyShoulder)</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>Materialise</td>
<td>0.979</td>
</tr>
<tr>
<td></td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>0.004</td>
</tr>
<tr>
<td>Medacta (MyShoulder)</td>
<td>Materialise</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>0.003</td>
</tr>
<tr>
<td>Materialise</td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.752</td>
</tr>
<tr>
<td></td>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>0.013</td>
</tr>
<tr>
<td>Zimmer-Biomet (PSI Shoulder System)</td>
<td>Tornier-Wright (Blueprint)</td>
<td>0.073</td>
</tr>
</tbody>
</table>
4. Discussion

The findings of this study did not support the hypothesis that the different methods would produce measurements that do not statistically differ. The biggest differences were observed in the inclination measurements, as the means ranged from 0,38° to 16,64°. The mean of 16,64° belongs to a method used by the investigators. Only taking the commercially available planning systems in consideration, the inclination means still ranged from 0,38° to 10,31°. The differences in version measurements were more subtle, yet sometimes still statistically significant. The version means ranged from -3,54° to 0,54°.

These results can have serious implications when it comes to pre-operative planning. As many studies have shown, excessive inclination and/or retroversion in TSA results in more frequent clinical failure. Likewise in RSA, improper positioning of the glenoid components could result in worse outcome. The fact that the inclination means range between 0,38° and 10,31° in commercially available systems is a cause for concern. Despite the fact that 3D measurements have proven to be more accurate than 2D measurements, we can still observe a significant difference due to the different methods that are used.

As with all newly emerging technologies, it might be necessary to implement a worldwide standard to perform these measurements that is both easily reproducible and clinically relevant. In the meantime, it is important to take a closer look at why there is such a difference between the methods used.

A. Intra- and interobserver reliability

The intra- and interobserver reliability were excellent for almost all measurements performed. Interobserver reliability for inclination was good for the IG-plane method (0.871) and excellent for the SM-plane method (0.952). The slightly lower interobserver reliability for the IG-plane method can be explained by multiple factors. To determine the IG-plane, all points of the antero-inferior glenoid rim (in some cases extended to also include the posterior rim) were marked with a software function. All these points were concentrated in a relatively small area, compared to the entire glenoid rim used in the SM-plane method. Furthermore, it became apparent that 3D reconstructed CT-scans sometimes showed small artefacts on the glenoid. This is a result of the Hounsfield Unit (HU) threshold that was chosen while performing the 3D reconstructions. These small artefacts don’t have a lot of influence on the measurements when you take the entire glenoid rim into consideration (in the SM-plane method) as the entire glenoid rim is large enough.
to cancel out these irregularities when you construct a mathematically best fitting circle. In the IG-plane method, these artefacts can have a bigger influence on the mathematically best fitting circle, as the antero-inferior rim is a relatively small surface area. A correction for this problem was performed by extending the marked area to include the posterior glenoid rim, when necessary. Nevertheless, this method remains very reproducible.

The interobserver reliability for version was excellent for both the IG-plane method (0.930) and the SM-plane method (0.983). The explanation for the slightly lower interobserver reliability of the IG-plane method is completely equivalent.

B. Inclination

The results of the post hoc Wilcoxon tests for inclination measurements show a statistically significant difference between:

1) IG-plane vs. all other methods
2) Zimmer-Biomet vs. all other methods
3) Materialise vs. Tornier-Wright

The biggest culprits for the observed differences are the IG-plane method and the method used by Zimmer-Biomet. When there is a significant difference in inclination measurements, the differences can be due to a difference in the transverse axis, a difference in the glenoid plane, or both. Both methods use the same transverse axis for the inclination measurements (glenoid centre – trigonum scapulae). Thus, the observed differences can best be explained by the different methods used to determine the glenoid plane. The differences between Materialise and Tornier-Wright have to be explained differently, as they use the same method to construct the glenoid plane.

B.1. IG-plane

The IG-plane had the largest mean inclination in this study at 16.64°. It is constructed on the antero-inferior rim of the glenoid with a best fit circle. This results in a plane that isn’t influenced by the prominent bony structure that is the tuberculum superior. All other methods in this study take the tuberculum superior into account when constructing the glenoid plane. The concave shape of the glenoid has as a result that a plane constructed on the inferior half of the glenoid, naturally has more inclination than a plane constructed on the entire glenoid (Fig III). Adding the
prominence of the tuberculum superior, it quickly becomes clear why the observed differences in inclination are so large.

![Diagram showing the difference in inclination between a plane constructed on the inferior half and a plane constructed on the entire glenoid.](image)

**Fig. III:** when constructing a plane on a concave glenoid, there is an observable difference of \( x^\circ \) in inclination between a plane constructed on the inferior half and a plane constructed on the entire glenoid. (simplified)

A second contributing factor to the differences observed, is a direct result of the methods used to construct the glenoid plane. As the tuberculum superior is left out of consideration, the centre of the glenoid circle will naturally be located a bit lower compared to the methods using the entire glenoid. The inclination measurements are performed using the transverse axis as a reference, and this axis is partly determined by the position of the glenoid centre.

A lower glenoid centre will tilt the transverse axis in such a way that the angle between the transverse axis and the glenoid plane decreases. Thus, the measured inclination will increase.

**B.2. Zimmer-Biomet**

Zimmer-Biomet (PSI Shoulder System) inclination was the lowest of all the methods in this study, with a mean of 0.38\(^\circ\). They provided the researchers with screenshots of every shoulder they performed measurements on, showcasing the method they used. 3 points are used to construct
the glenoid plane: an anterior, posterior and superior point on the glenoid. While observing these screenshots, it became clear that their method is the only one in which a part of the inferior glenoid rim was left out of consideration. The inferior rim however, can often times be quite prominent, much like the tuberculum superior, which Zimmer-Biomet does take in consideration.

Fig IV: screenshots provided by Zimmer-Biomet showcasing the reference points used for their measurements on the glenoid (A and B). When drawing a simplified plane through the 3 points on the glenoid rim (C), the inferior rim remains exposed and the plane has less inclination than the methods using the entire glenoid (or the inferior half of the glenoid).

Fig. V: when constructing a plane on a concave glenoid, there is an observable difference of $\gamma^*$ between the methods used by Zimmer-Biomet and the methods using the entire glenoid. (simplified)

Similar to what was observed in the IG-plane method, we can describe a second contributing factor to the observed differences. As the inferior rim is left out of consideration, the mathematical
centre of the glenoid will naturally be located a bit higher compared to the methods using the entire glenoid. The inclination measurements are performed using the transverse axis as a reference, and this axis is partly determined by the position of the glenoid centre. A higher glenoid centre will tilt the transverse axis in such a way that the angle between the transverse axis and the glenoid plane increases. Thus, the measured inclination will decrease.

**B.3. Materialise and Tornier-Wright**

Materialise and Tornier-Wright (Blueprint) were the only pair that showed statistically significant differences in inclination measurements, apart from pairs with Zimmer-Biomet or the IG-plane. Materialise had a mean inclination of 5.38°, while Blueprint had a mean inclination of 10.31°. This is unusual, as both companies use the same method to determine the glenoid plane and the reference transverse axis. The only difference between the two companies is that Blueprint is entirely automated. As other studies have suggested, the automatic 3D reconstruction of the glenoid can sometimes be inaccurate. The humeral head needs to be substracted from the CT scans in order to reconstruct the glenoid. Often times the humeral head is not completely substracted from the glenoid fossa, leaving extra bony fragments on the glenoid surface, which can in turn influence the measurements. Conversely, the humeral head substraction can be excessive, resulting in gaps in the glenoid surface. In this study however, only 1 of 13 glenoids was found to be reconstructed improperly. The differences observed between Materialise and Blueprint might be due to the proprietary algorithms used to determine the mathematical best fitting glenoid plane and/or the transverse axis.

**C. Version**

After comparing all the different methods in a pairwise manner, only 4 significantly different pairs were found. Only one of these pairs (SM-plane and Zimmer-Biomet) did not include Medacta. The significant difference between SM-plane and Zimmer-Biomet may be explained by the points used in the determination of the glenoid plane. The method used by Zimmer-Biomet does not include a point on the inferior glenoid rim. This means that the glenoid plane determined by Zimmer, does not take into account the angulation of the inferior rim. This may influence the measurements of glenoid version.

All other significant results of the pairwise comparisons for version involved Medacta. Medacta differed significantly from IG-plane (p=0.006), Materialise (p=0.002) and Zimmer-Biomet
(p=0.003). As IG-plane, Zimmer-Biomet and Medacta use the same 3 points to define the scapular plane (glenoid centre, trigonum scapulae, inferior scapular point), it can be assumed that the explanation for these differences must lie in the way that the glenoid centre is determined. The method used by Medacta is unclear, but states that the glenoid centre corresponds to the deepest or most medial point of the glenoid fossa. In order for this method to be reproducible, a certain consistency is required regarding the position of this most medial point in the glenoid fossa. One theory states that there might be a tendency for these medial points to be positioned towards the anterior part of the glenoid, resulting in higher retroversion measurements. However, no literature on the subject of glenoid fossa morphology was found to support this claim.

**D. Clinical Relevance**

Given the clinical consequences of these measurements, it is equally important to look at the clinical relevance of the different methods. A very important factor in this matter, is the eventual shoulder arthroplasty performed, based on these measurements. During the procedure, the inferior part of the glenoid is used as the main anchor point for the prosthesis. Thus, it would seem logical to use methods that take into account the inferior rim of the glenoid. As a consequence, the method used by Zimmer-Biomet might be less suitable for pre-operative 3D planning, whereas the IG-plane method would be more appropriate. However, literature shows that this may vary depending on the performed procedure (TSA or RSA).

**5. Conclusion**

This study shows that the wide variety of methods used by different companies and institutions, results in significantly differing measurements for glenoid inclination and version. The most notable differences were found in the measurement of inclination. These differences were less spectacular in the comparisons of version measurements, but still present. Given the major consequences these measurements have on the eventual arthroplasty, it is clear that a more refined and universally used method is required. Literature supports the claim that this method should be an inferior glenoid based method, given the clinical relevance of the inferior glenoid. Generally, it is clear that more research is required on this matter.
6. References