Digital Planning Software Fails to Reflect Stem Torsion on Plain Radiographs after Total Hip Arthroplasty

Digitale Planungssoftware ist nicht geeignet die Torsion des Implantatschaftes nach endoprothetischem Hüftersatz zu bestimmen

Authors M. Worlicek[1], M. Weber[1], F. Zeman[2], M. Wörner[1], M. Schneider[1], J. Grifka[1], T. Renkawitz[1], B. Craiovan[1]

Affiliations
[1] Departement of Orthopedic Surgery, University of Regensburg, Bad Abbach, Germany
[2] Center of Clinical Studies, University of Regensburg, Regensburg, Germany

Key words
- hip arthroplasty
- stem version
- anterior posterior radiographs
- planning software
- CT

Zusammenfassung

Zielsetzung: Ziel dieser Studie war es die Validität einer kommerziell verfügbaren Planungssoftware zur Vermessung von zweidimensionalen Röntgenbildern, im Vergleich zu CT-Aufnahmen zu bestimmen.
Patienten und Methoden: In einer sekundären Analyse einer großen prospektiven, kontrollierten Studie (DRKS00 000 739, Deutsches Register klinischer Studien 02.05.2011), wurden anteriorposteriore Hüftübersichtsaufnahmen und 3D-CT-Aufnahmen einer Untergruppe von 121 Patienten, die einen endoprothetischen Hüftersatz erhielten, untersucht. Der Eingriff wurde über einen minimalinvasiven Zugang, in Seitenlage, unter Verwendung eines nichtzementierten, kegelförmigen Geradschaftes durchgeführt. Zur Messung der Schafttorsion verwendeten wir eine digitale Planungssoftware (TraumaCad 2.0, BrainLAB Feldkirchen, Deutschland). Die Messungen wurden von zwei unabhängigen Untersuchern durchgeführt und nach sechs Wochen wiederholt. Die Ergebnisse wurden mit 3D-CT-Messungen eines unabhängigen, verblindeten, externen Instituts verglichen. Die Untersuchung wurde durch die ansässige Ethikkommission (Nr. 10-121-0263) genehmigt.
Ergebinsse: Die Messungen der Röntgenbilder zeigten eine sehr hohe Übereinstimmung, sowohl zwischen den Untersuchern, als auch zwischen den Erst- und Zweitmessungen der einzelnen Untersucher. Die Intra-Klassen-Korrelation der Erst- und Zweitmessungen lag bei 0,97 für Untersucher 1 und bei 0,98 für Untersucher 2. Die Zuverlässigkeit der Messungen zwischen beiden Untersuchern lag bei 0,99 bzgl. der Durchschnittswerte beider Messungen. Die mittlere Abweichung zwischen den Durchschnittswerten der Messungen auf den Röntgenbildern und den 3D-CT Messungen lag bei 0,41° (SA 11,24°) (Messbereich −33,85°–22,50°; 95 % limits of agreement −21,63 – 22,45), allerdings

Abstract

Purpose: The purpose of this study was to evaluate the validity of commercially available planning software on plain radiographs after THA compared to CT scans as the gold standard.
Patients and Methods: In a prospective clinical study, anteroposterior (AP) radiographs and three-dimensional CT scans (3D-CT) were obtained for 121 patients, who underwent minimally invasive, cementless THA with a straight tapered stem, in a lateral decubitus position. For measuring SV, we used digital planning software (TraumaCad 2.0, BrainLAB Feldkirchen, Germany). Two independent raters repeated the analysis after a six-week interval. Radiological measurements were compared with 3D-CT measurements by an independent, blinded external institute. This investigation was approved by the local ethics commission (no. 10-121-0263) and is a secondary analysis of a larger project (DRKS00 000 739, German Clinical Trials Register May-02 – 2011).
Results: The radiograph measurements showed very high intra- and interrater agreement. The intra-class correlation (ICC) of the intrarater agreement was 0.97 for rater 1 and 0.98 for rater 2. The intrarater reliability was 0.99 using the mean values of both rater measurements. The mean difference between the average radiograph measurement and the 3D-CT-based measurement was 0.41° (SD 11.24°) (range: −33.85°–22.50°; 95 % limits of agreement: −21.63 – 22.45), but there was no correlation found between both methods.
Conclusion: Measuring stem version with the help of commercially available digital planning software on plain radiographs after THA has high intra- and interrater reliability but clinically inacceptable validity and reliability when compared to 3D-CT scans.

received 20.2.2016
accepted 13.4.2016

Bibliography
DOI http://dx.doi.org/10.1055/s-0042-106975
Published online: 29.6.2016
Fortschr Röntgenstr 2016; 188: 763–767 © Georg Thieme Verlag KG Stuttgart · New York · ISSN 1438-9029

Correspondence
Dr. Michael Worlicek
Departement of Orthopedic Surgery, University of Regensburg
Kaiser-Karl V. Allee 3
93077 Bad Abbach
Germany
Tel.: ++ 49/94 05/18 48 32
Fax: ++ 49/94 05/18 24 79
michael.worlicek@gmx.de

konnte keine Korrelation zwischen beiden Methoden nachgewiesen werden.

Schlussfolgerung: Die Vermessung der Schafttorsion mit einer kommerziell verfügbaren, digitalen Planungssoftware auf zweidimensionalen Röntgenbildern, nach endoprothetischem Hüftersatz, zeigt trotz guter Reproduzierbarkeit, eine klinisch inakzeptable Ungenauigkeit verglichen mit 3D-CT-Aufnahmen.

Kernaussagen:

▶ Die Vermessung der Schafttorsion auf zweidimensionalen Röntgenbildern mit digitaler Planungssoftware ist nicht valide.

Key Points:

▶ Measuring stem torsion after THA on plain radiographs with digital planning software is not valid.

Citation Format:

▶ Worlicek M, Weber M, Zeman F et al. Digital Planning Software Fails to Reflect Stem Torsion on Plain Radiographs after Total Hip Arthroplasty. Fortschr Röntgenstr 2016; 188: 763–767

Introduction

Primary total hip arthroplasty (THA) is one of the most performed orthopedic operations worldwide [1]. However, inaccurate placement of the femoral and acetabular components in THA can lead to dislocation, decreased range of motion (ROM), periprosthetic or bony impingement and component wear [2 – 6]. If a complication in THA appears, it is necessary to detect the reason and whether wrong component placement could be responsible. Therefore, the surgeon is able to resolve the problem selectively. The easiest way would be to analyze standard radiographs. Previous studies showed that version of acetabular component can be measured on single anteroposterior (AP) radiographs with accurate results considering the measuring method for daily clinical use [7 – 9]. The stem version can be evaluated using the so-called "Budin method", a validated protocol for the radiological measurement of stem version [10]. A limitation of this method is the need for a special radiological image, which leads to additional radiation and expense. Another problem regarding exact assessment of the femoral component is its variation in the final position. Sendtner et al. found a range from –19° retroversion to 33° anteversion in cementless THA, which is in accordance with the results of Wines et al., who showed a range from –15°–52° [11, 12]. Weber et al. developed a new mathematical formula for measuring stem version using the projected prosthetic neck-shaft angle (NSA) on AP radiographs and compared the results with three-dimensional CT scans (3D-CT) [13]. The authors found a high reliability and validity in the evaluation of the stem version in cementless THA, considering the limitation that this method cannot differentiate between anteversion and retroversion.

Today there are several software programs that are used for preoperative planning in THA and total knee arthroplasty (TKA). Some of these programs offer the possibility of postoperative component measuring in AP radiographs. This can be relevant considering claims of recourse of unsatisfied patients. However, to the best of our knowledge, no study has been reported about the validity of these programs with respect to measuring the stem version (SV) after THA.

This retrospective secondary analysis out of a large prospective study aimed to investigate the objectivity, reliability and validity of measuring stem version with the help of commercially available planning software on plain radiographs after THA when compared to CT scans as the gold standard.

Patients and Methods

In the course of a registered, prospective controlled trial (DRKS00 000 739, German Clinical Trials Register), hip radiographs in two planes (AP and axial) and 3D-CT scans were obtained for patients who underwent minimally invasive THA. This investigation was approved by the local ethics committee (no. 10-121-0263). All procedures were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2000. The current study is a secondary analysis of a larger project [14].

The primary outcome of this larger study was to assess whether the ROM of the prosthetic joint could be improved by computer-assisted functional optimization of the position and containment of the acetabular component.

For this study AP radiographs of 121 patients out of the whole study collective were analyzed. Characteristics of the study group are shown in **Table 1**. The patients were chosen by random. All THAs were performed by four orthopedic surgeons (JG, ES, MWö, TR) from Regensburg University Medical Center. Each surgeon had experience with > 200 fluoroscopy and navigation-controlled THAs. Press-fit acetabular components, uncemented hydroxyapatite-coated femoral components (Pinnacle acetabular component, Corail femoral component (both DePuy, Warsaw, Indiana), neutral polyethylene liners and metal heads with a diameter of 32 mm were used in all patients.

All operations were performed in the lateral decubitus position through a minimally invasive, modified Smith-Petersen approach (MicroHip®) [15]. Six week after surgery, full weight-bearing standing radiographs of the whole pelvis AP and the operated hip axial were taken (MULTIX TOP ACSS; Siemens, Erlangen, Germany). The radiographers made sure that the pelvis was set parallel to the plane of the film without rotation or flexion of the hip joint and the leg was placed in a neutral position, with the patella pointing forward disregarding the foot progression angle in the event of a tibial version. All radiographs were to be taken under these standardized conditions (focus-film distance 115 cm, 75 kV, automatic exposure). During the same visit a CT scan was obtained from the pelvis down to the femoral condyles (Somatom Sensation 16; Siemens, Erlangen, Germany).

The radiological SV was measured with the help of the "semi-automatic" function of digital planning software (TraumaCad 2.0, BrainLAB Feldkirchen, Germany). For this purpose, an exact circle has to be drawn around the femoral head to assess its center. Then the axes of the component neck and the axis of the stem have to be determined. The angle between these axes is regarded as the neck-shaft angle measured automatically by the software or manually with a 4-point middle line. Both anteversion and retroversion around the axis of the femur cause the projected neck-shaft to appear increased (**Fig. 1**). The true NSA of the stem is known to be 135° and its difference to the vertical axis of the implant (180°) is 45°. This means the higher the version of the stem, the higher the projection-based increase of the NSA. The software promises to recognize the difference between the neck-

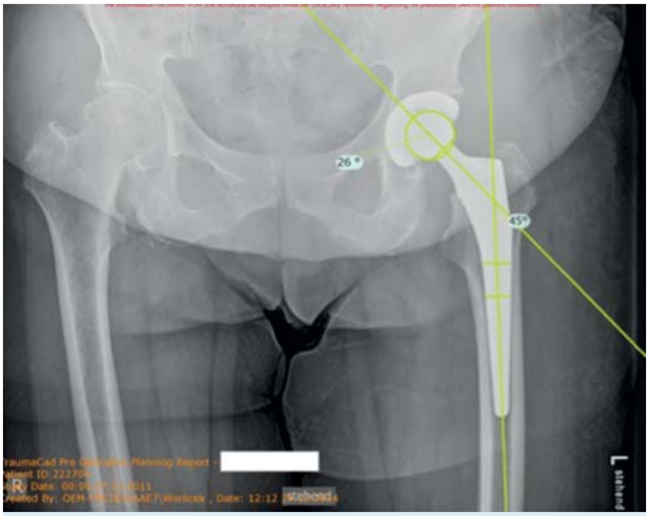

Fig. 1 Radiographic assessment of stem version (26°) calculated automatically by the software, the 45° are the difference between the true neck-shaft angle of 135° of the implant and the vertical stem axis (180°).

Abb. 1 Vermessung der Schafttorsion (26°), automatisch errechnet durch die Software, die 45° sind die Differenz zwischen dem wahren Hals-Schaftwinkel von 135° und der vertikalen Schaftachse (180°).

Table 1 Characteristics of the study group.[1]

Tab. 1 Charakterisitka des Studienkollektivs.

	n = 121
gender (female) (%)	66(55)
age (yrs)	62.7 (SD 0.6)
BMI (kg/m^2)	27.0 (SD 4.1)
treatment side (right) (%)	67 (55)
femoral component size (IQR)	11 (2)
femoral component geometry (%)	Std 60 (49.5). HO 61 (51.5)
OP time (min)	71.4 (12.5)
kellgren (IQR)	8 (1)
length of incision (cm) (SD)	10. 4 (1.3)

BMI: body mass index; HO: high-offset stem; Std: standard stem; IQR: interquartile range.
[1] For categorical data, values are given as relative and absolute frequencies; for quantitative data, values are given as mean with SD in parentheses.

shaft angle and the projection-based angle and converts it into the degree of version of the stem. The software is not able to differentiate between anteversion and retroversion. For that reason the second clinical standard plane (axial) has to be consulted. All radiological measurements were performed by two independent examiners (MW, MS), who repeated the measurements after a six-week interval. The raters were blinded to the 3D-CT values as well as to each other's results. In addition, 3D-CT assessment of prosthetic stem version was obtained by an independent, blinded external institute (MeVis Medical Solutions, Bremen, Germany), as described by Sendtner et al. [11]. Correlation was characterized as poor (0.00 – 0.20); fair (0.21 – 0.40); moderate (0.41 – 0.60); good (0.61 – 0.80) or excellent (0.81 – 1.00) [16]. As the generally accepted range of stem anteversion is between 10° and 20°, we defined a tolerance limit of 5° compared with 3D-CT as clinically acceptable [17].

Statistical methods

Statistical analyses were performed with IBM SPSS Statistics® 23.0 (SPSS Inc., Chicago, IL, USA) and R version 3.2.1. Data are presented as mean, standard deviation and range. The accuracy of the radiographs was assessed using Bland-Altman plots and clinical evaluation. Bland-Altman plots illustrate the accuracy of the radiograph measurements compared to the 3D-CT-based measurement (gold standard) by plotting the gold standard on the x-axis and the difference of both measurements on the y-axis. The dashed lines in the graph represent the 95 % limits of agreement (mean ± 1.96 SD). Intra- and inter-rater agreement (precision) was assessed by the intraclass correlation coefficient (ICC) and graphically by scatter plots and standard Bland-Altman plots (i. e., the x-axis shows the average of both measurements).

Results

Precision

The radiograph measurements showed very high intra- and interrater agreement. The ICC of the intrarater agreement was 0.97 for rater 1 and 0.98 for rater 2 (**Fig. 2**). The interrater reliability was 0.99 using the mean values of both rater measurements. The 95 % limits of agreement range between –7.4 and 6.6 for the intrarater agreement and between –6.9 and 6.2 for the interrater agreement. The mean difference is close to null in both cases (**Fig. 2**).

Accuracy

Due to the excellent intra- and interrater agreement, we used the mean value of the four measurements for the Bland-Altman plot. The mean difference between the average radiograph measurement and the 3D-CT-based measurement was 0.41° (SD 11.24°) (range: –33.85°–22.50°; 95 % limits of agreement: –21.63 – 22.45) (**Fig. 3**). In all, 43/121 (36 %) of the radiological measurements of prosthetic SV were within a tolerance limit of 5° compared with 3D-CT.
The Bland-Altman plot shows that there was no systematic error of the radiograph measurements. **Table 2** summarizes the measurements on plain radiographs performed by the two raters and by 3D-CT.

Discussion

Malpositioning of components in THA leads to pain, reduced range of motion and early instability [2 – 5, 18]. So far, CT has been the gold standard for postoperative assessment of THA components because of its high accuracy and reliability [10, 19]. Today several software programs promise the ability to measure SV on standard AP radiographs. We aimed to investigate the objectivity, reliability and validity of measuring stem version with the help of commercially available planning software on plain radiographs after THA compared to CT scans as the gold standard. We found excellent intra- and interrater reliability of the software. The raters even had the same outliers in both their measurements compared to 3D-CT.
In regard to the software's reliability and validity, the mean prosthetic version of the stem measured by both raters was close to the mean version measured by 3D-CT but without correlation between the two techniques. Reasons for inaccuracy could be differences in picking landmarks, ignorance of the femoral tilt, the angular difference between the long axis of the femoral stem

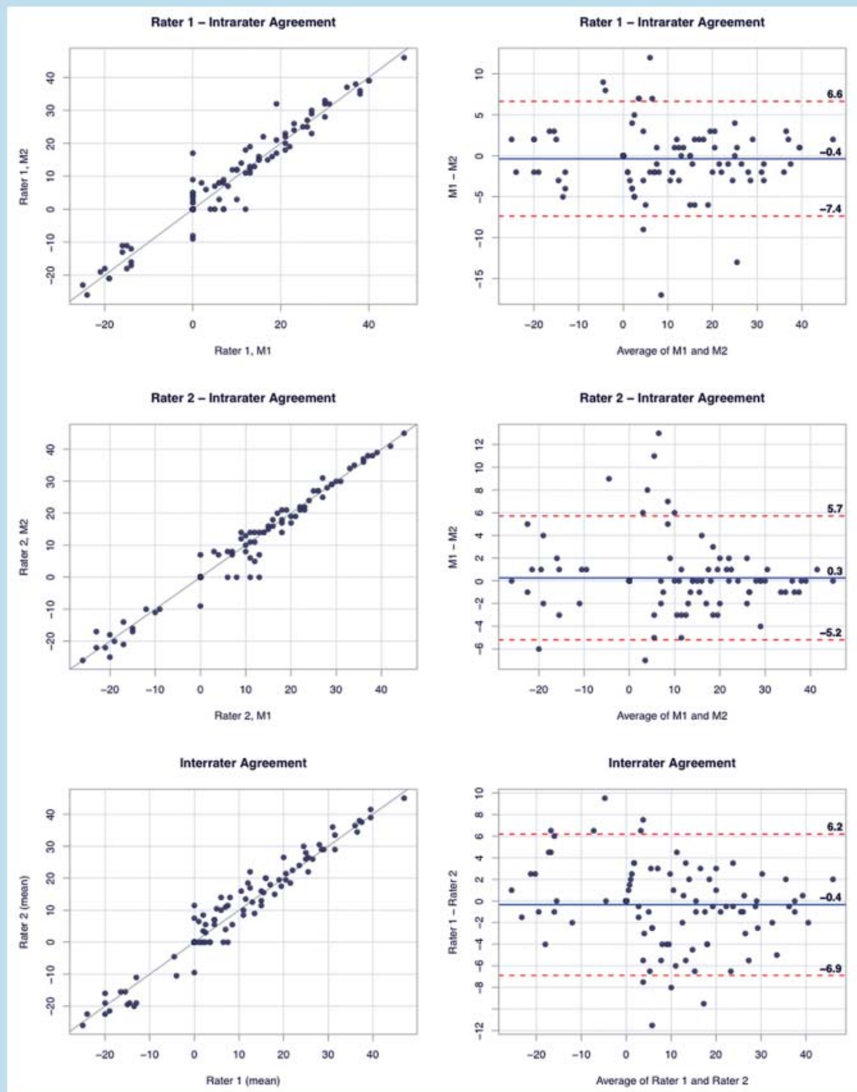

Fig. 2 The Scatterplots and Bland-Altman plots show a high precision within both measurements of both observers and between both observers (anteversion (°)).

Abb. 2 Die Scatterplots und die Bland-Altman-Plots zeigen die hohe Präzision beider Messungen der Untersucher und zwischen den Messungen der beiden Untersuchern (Antetorsion (°)).

Table 2 Durchschnittswerte der Messungen beider Untersucher auf den Röntgenbildern im Vergleich zu den 3D-CT Daten.

Tab. 2 Mean values of the measurements of both raters on plain radiographs compared to the 3D-CT data.

	rater 1 M 1	rater 1 M 2	rater 2 M 1	rater 2 M 2	average rater 1	average rater 2	average raters	3D-CT
mean	7.1	7.4	7.7	7.7	7.2	7.6	7.4	7.8
SD	14.5	14.8	15.2	15.3	14.5	15.3	14.8	9.2
minimum	–25.0	–26.0	–26.0	–26.0	–25.0	–26.0	–25.5	–18.9
maximum	48.0	46.0	45.0	45.0	47.0	45.0	46.0	37.7
ICC	0.97		0.98		0.99			

SD = standard deviation, M = measurement, ICC = intraclass correlation coefficient.

and the mechanical axis on a sagittal radiograph. That difference is due to the fact that the stem of the prosthesis follows the natural anterior bow of the proximal femur [20]. In summary, we found high reliability but no validity for the use of digital planning software for measuring SV after THA. Nevertheless, we think the software can be a useful tool for a first approximate determination of major rotational errors of the femoral component in a painful hip after THA.

There are several limitations when measuring SV with the help of digital planning software on plain radiographs. First, software is not able to differentiate between anteversion and retroversion. Therefore, a second axial radiograph is needed to distinguish between the two. Second, we found the handling of the software itself challenging. Nevertheless, we found excellent intra- and inter-rater reliability. The exact determination of the axis of the neck and the stem is prone to error, because it must be done by hand so landmark selection is inaccurate. Even minimal changes of the position of one axis lead to a high change of the value of SV. A more accurate way for determination of the axes could be to draw concentric circles into the neck and the stem and use their midpoints as orientation for the axes as described by Weber et al. [21]. Third, the software bases its measurements on the known NSA. A factor that has a high influence on the NSA is the position of the patient. This means any internal or external rotation of the leg and any extension or flexion of the hip can lead to a misinterpretation of the degree of the stem version. To avoid that impact, a rigorously

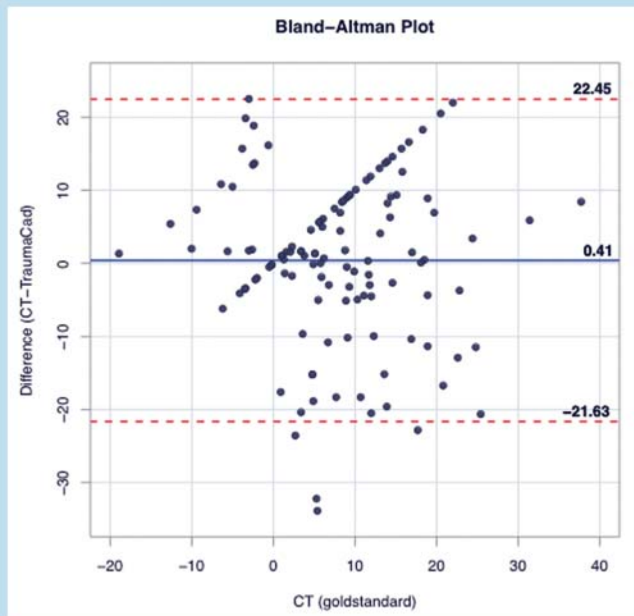

Fig. 3 The Bland-Altman plot shows the mean difference between the average of the measurements of both observers and 3D-CT measurements of stem torsion. The straight line represents the mean value of all differences between the pair of measurements and the dashed lines above and below represent 95 % limits of agreement (anteversion (°)).

Abb. 3 Der Bland-Altman-Plot zeigt die mittlere Abweichung zwischen den Durchschnittswerten der Messungen beider Untersucher und der 3D-CT-Messung der Schafttorsion. Die gerade Linie repräsentiert den Mittelwert der Differenzwerte der beiden Messungen und die gestrichelten Linien begrenzen das 95 % Übereinstimmungsintervall (Antetorsion(°)).

standardized radiological technique as used by us in the study protocol is essential to ensure exact positioning of the patient and to minimalize projection errors [22]. Furthermore, the quality of the radiograph is very important. Although our measurements were obtained under optimized conditions following a strict protocol, the radiological image in two cases was inadequate. So there might be another limitation for the method in clinical practice. A different option for measuring SV after THA on plain radiographs is the “Budin method”. In short, this method uses a posteroanterior radiograph of the hip in 90° flexion and 30° abduction and with 90° flexion of the knee [10]. Lee et al. found a high correlation between the radiological and CT measurements ($r = 0.88$, $p < 0.001$) with excellent intra- (0.94) and interrater reliability (0.93) [10]. Another accurate method for the assessment of SV is to use a mathematical formula, which puts the projected NSA of the stem on plain AP radiographs in relation to the true NSA of the implant. This method was described by Weber et al. and was valid compared to 3D-CT ($r = 0.88$, $p < 0.001$) [13].

Conclusion

In conclusion, measuring stem version with the help of commercially available digital planning software on plain radiographs after THA has high intra- and interrater reliability but clinically inacceptable validity and reliability when compared to 3D-CT scans.

References

1 *Learmonth ID, Young C, Rorabeck C.* The operation of the century: total hip replacement. Lancet 2007; 370: 1508 – 1519

2 *D'Lima DD, Urquhart AG, Buehler KO et al.* The effect of the orientation of the acetabular and femoral components on the range of motion of the hip at different head-neck ratios. The Journal of bone and joint surgery American volume 2000; 82: 315 – 321

3 *Lewinnek GE, Lewis JL, Tarr R et al.* Dislocations after total hip-replacement arthroplasties. The Journal of bone and joint surgery American volume 1978; 60: 217 – 220

4 *Malik A, Maheshwari A, Dorr LD.* Impingement with total hip replacement. The Journal of bone and joint surgery American volume 2007; 89: 1832 – 1842

5 *Patel AB, Wagle RR, Usrey MM et al.* Guidelines for implant placement to minimize impingement during activities of daily living after total hip arthroplasty. The Journal of arthroplasty 2010; 25: 1275 – 1281 e1271

6 *Renkawitz T, Worner M, Sendtner E et al.* Principles and new concepts in computer-navigated total hip arthroplasty. Der Orthopade 2011; 40: 1095 – 1102

7 *Craiovan B, Renkawitz T, Weber M et al.* Is the acetabular cup orientation after total hip arthroplasty on a two dimension or three dimension model accurate? International orthopaedics 2014; 38: 2009 – 2015

8 *Goergen TG, Resnick D.* Evaluation of acetabular anteversion following total hip arthroplasty: necessity of proper centring. The British journal of radiology 1975; 48: 259 – 260

9 *Miller TT.* Imaging of hip arthroplasty. European journal of radiology 2012; 81: 3802 – 3812

10 *Lee YK, Kim TY, Ha YC et al.* Radiological measurement of femoral stem version using a modified Budin method. The bone & joint journal 2013; 95-B: 877 – 880

11 *Sendtner E, Tibor S, Winkler R et al.* Stem torsion in total hip replacement. Acta orthopaedica 2010; 81: 579 – 582

12 *Wines AP, McNicol D.* Computed tomography measurement of the accuracy of component version in total hip arthroplasty. The Journal of arthroplasty 2006; 21: 696 – 701

13 *Weber M, Lechler P, von Kunow F et al.* The validity of a novel radiological method for measuring femoral stem version on anteroposterior radiographs of the hip after total hip arthroplasty. The bone & joint journal 2015; 97-B: 306 – 311

14 *Renkawitz T, Haimerl M, Dohmen L et al.* Minimally invasive computer-navigated total hip arthroplasty, following the concept of femur first and combined anteversion: design of a blinded randomized controlled trial. BMC musculoskeletal disorders 2011; 12: 192

15 *Michel MC, Witschger P.* MicroHip: a minimally invasive procedure for total hip replacement surgery using a modified Smith-Peterson approach. Ortopedia, traumatologia, rehabilitacja 2007; 9: 46 – 51

16 *Landis JR, Koch GG.* The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159 – 174

17 *Dorr LD, Wan Z, Malik A et al.* A comparison of surgeon estimation and computed tomographic measurement of femoral component anteversion in cementless total hip arthroplasty. The Journal of bone and joint surgery American volume 2009; 91: 2598 – 2604

18 *Pierchon F, Pasquier G, Cotten A et al.* Causes of dislocation of total hip arthroplasty. CT study of component alignment. The Journal of bone and joint surgery British volume 1994; 76: 45 – 48

19 *Murphy SB, Simon SR, Kijewski PK et al.* Femoral anteversion. The Journal of bone and joint surgery American volume 1987; 69: 1169 – 1176

20 *Renkawitz T, Haimerl M, Dohmen L et al.* The association between Femoral Tilt and impingement-free range-of-motion in total hip arthroplasty. BMC musculoskeletal disorders 2012; 13: 65

21 *Weber M, Woerner ML, Springorum HR et al.* Plain radiographs fail to reflect femoral offset in total hip arthroplasty. The Journal of arthroplasty 2014; 29: 1661 – 1665

22 *Merle C, Waldstein W, Pegg EC et al.* Prediction of three-dimensional femoral offset from AP pelvis radiographs in primary hip osteoarthritis. European journal of radiology 2013; 82: 1278 – 1285