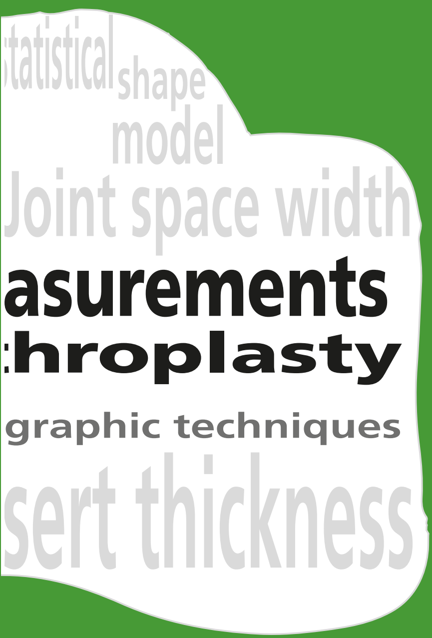
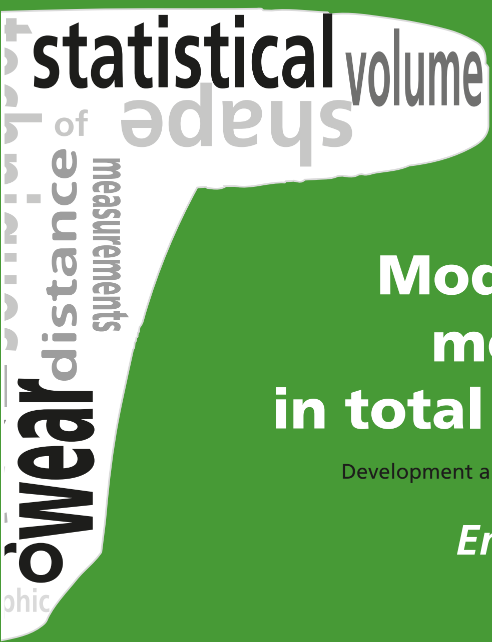
Author: IJsseldijk, E.A. van
Title: Model-based wear measurements in total knee arthroplasty : development and validation of novel radiographic techniques
Issue Date: 2016-09-01





Model-based wear measurements in total knee arthroplasty

Development and validation of novel radiographic techniques

Emiel van IJsseldijk

Model-based wear measurements in total knee arthroplasty

Development and validation of novel radiographic techniques

Emiel van IJsseldijk

Model-based wear measurements in total knee arthroplasty
– *Development and validation of novel radiographic techniques*
E.A. van IJsseldijk
PhD thesis, Leiden University Medical Center, Leiden, The Netherlands

Financial support for this thesis was given by Anna fonds

ISBN 978-90-9029866-5
Design: Grafisch Vormgever Jaap van der Heijden
Printed by: GVO drukkers & vormgevers B.V.

Model-based wear measurements in total knee arthroplasty

Development and validation of novel radiographic techniques

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.M.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 1 september 2016
klokke 16:15

door

Emiel Arno van IJsseldijk
geboren te Schiedam
in 1981

Promotor

Prof.dr.ir. E.R. Valstar

Copromotores

Dr.ir. B.L. Kaptein

Dr. B.C. Stoel

Leden promotiecommissie

Prof.dr R.G.H.H. Nelissen

Prof.dr.ir. M.A. Viergever, Universitair Medisch Centrum Utrecht

Prof.dr. M.G. Kloppenburg

Dr. M. Harman, Clemson University, USA

Inhoud

1

Introduction

Introduction

1-1 Osteoarthritis of the knee

Osteoarthritis (OA) is a multi-factorial joint disease characterized by progressive degeneration of cartilage tissue thickening of the joint's subchondral bone resulting in a painful and stiff joint with decreased limited mobility [1, 2]. Osteoarthritis is a disease on its own, but can also occur secondary to an inflammatory disease like rheumatoid arthritis, post-traumatic or after congenital or acquired limb deformities.

In 2011, approximately 594,000 Dutch inhabitants suffered from knee OA, which is approximately 4% of the total population [3]. Knee osteoarthritis affects especially elderly patients. In the Netherlands, the registered prevalence of this disease for patients over the age of 65 years was 6.4% and 11.2% for men and women respectively [3]. This is a major health care burden, resulting in 1.11 billion euros for the direct and indirect healthcare costs of osteoarthritis in the Netherlands alone in 2011. This is 1.2% of the total annual Dutch healthcare costs [4].

Due to an aging population and longer life expectancy the prevalence of knee osteoarthritis is increasing. Moreover, since obesity – which is a risk factor for osteoarthritis – is ever more present in our society, more and younger patients will be affected by osteoarthritis [5-9].

1-2 Total Knee Arthoplasty

Total Knee Arthroplasty (TKA) is an effective treatment for endstage symptomatic osteoarthritis [10, 11]. TKA is a surgical procedure in which the knee joint is replaced by a prosthesis consisting of a femoral component, a tibial component, a polyethylene insert facilitating the articulation between de femoral and tibial component and - in some cases - a patella component (Figure 1-1).

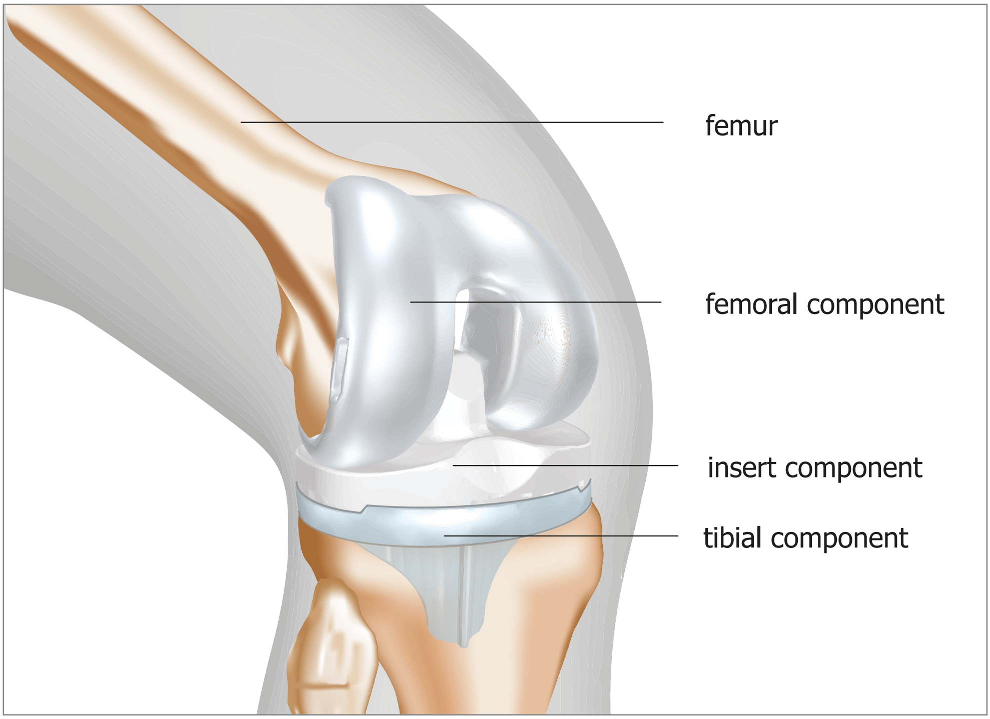


Figure 1-1. Illustration of a knee prosthesis and its components.

Performed widely from the 1970's, the procedure is well-established as a successful treatment in relieving pain and restoring joint function for patients with end stage osteoarthritis [11]. In general, an incision is made longitudinally across the knee, the joint capsule is opened and the patella (knee cap) is exposed by rotating it to the lateral side. Further dissections are made until the distal femur and proximal tibia are sufficiently exposed. The prosthesis' components are fixated to the corresponding bones. Before this can be done, bone has to be removed so that the artificial components have a close fit. Fixation of the components can be obtained either by applying bone-cement or by bone ingrowth.

Many different types and designs of knee prostheses are available, like posterior stabilized types, cruciate retaining types or rotating platform types. These types were developed to improve the kinematics or the stability of the prosthesis. Depending on the prosthesis type, the cruciate ligaments are resected or retained, the patella is resurfaced or not or additional collateral ligament balancing is required. As for rotating platform tibia components, this design allows for higher tibiofemoral conformity without undue kinematic constraint. Even though TKA is generally successful, implant failure remains a significant problem. National registries report a failure rate of 5% to 10% at 10 years after the initial surgical procedure, indicating revision surgery was required [9, 12, 13]. Revision surgeries are extensive procedures with higher intra- and post-operative risks. Moreover, revision TKAs have a higher risk of revision (i.e. re-revision) and lower patient satisfaction compared to primary TKAs [14]. Altogether, the impact of implant failure on patients and healthcare costs is substantial [12].

In line with the increase in the number of patients suffering from osteoarthritis of the knee, the incidence of TKAs is also expected to increase. Therefore, the impact of implant failure will increase as well. As implant failure is more frequent for younger, more active patients and the prevalence of TKAs for younger patients increases [14, 15], the impact of implant failure is aggravated further. In order to reduce patient consequences and the financial burden of TKA procedures, reduction of the number of implant failures is an important topic in both clinical and technical research.

1-3 Polyethylene wear

The insert of a TKA is made of ultrahigh molecular weight polyethylene, which is designed to withstand the sliding and rolling articulation of the femoral component in the daily use of the prosthesis. However, wear of this component occurs due to various factors such as e.g. excessive forces during articulation, poor quality of the polyethylene material and poor alignment of the implant components increasing the load on articulating surfaces [16-18].

Wear particles ranging in size between 0.1 microns to 0.5 millimeter are released in the wear process. Especially the smaller particles can cause a local inflammatory reaction, which is associated with bone resorption around the TKA resulting in osteolysis and eventually aseptic loosening of the prosthesis [19, 20]. Aseptic loosening is an important failure mechanism as it is related to one out of every four revisions [21, 22].

Besides, TKA failure can occur for severe wear cases when metal to metal contact between the prosthesis elements occurs, resulting in an irreversibly damaged and non-functional prosthesis.

1-4 Relevance of measuring wear

It has been shown that the rate at which the remaining insert thickness decreases can predict TKA failure [23, 24]. For this reason, an accurate and precise method is required to assess the progression of polyethylene wear *in vivo*, which can be used to predict (future) instability and loosening and thereby support clinical decision making as to initiate a timely intervention or to decide which patients should be monitored more intensively. Timing is very important to minimize the burden of surgery for both the patient and the surgeon [12]. On the one hand, the surgical procedure should not be performed too soon as to prevent unneeded risks for patients. On the other hand, postponing the revision surgery may lead to an inferior outcome in case of high wear rates, due to the progressing osteolysis (reducing the bone stock available) and the increased inflammation related to wear debris.

A second application for an accurate and precise wear measurement is to evaluate the wear resistance of (new) prosthesis designs [25]. Wear characteristics of new prosthesis designs are currently assessed with knee simulator studies before market introduction. These simulator studies apply repetitive loads and motion to the prosthesis based on models of daily patient activity, yet they are limited in incorporating patient-specific effects and events such as extreme usage or simple missteps [16, 26]. An accurate and precise measurement of polyethylene wear *in vivo* is therefore also required to monitor the quality of new and existing prosthesis designs.

1-5 Wear measurement techniques

In current clinical practice, weight-bearing planar radiographs are the clinical standard for the assessment of wear *in vivo*[27, 28]. In these images, the remaining polyethylene insert thickness is estimated using the minimum joint space width (mJSW) measurement, in which the apparent distance between the metal tibial tray and the femoral condyles is assessed[27, 29-31]. An example of an image with these reference objects is shown in Figure 1-2.

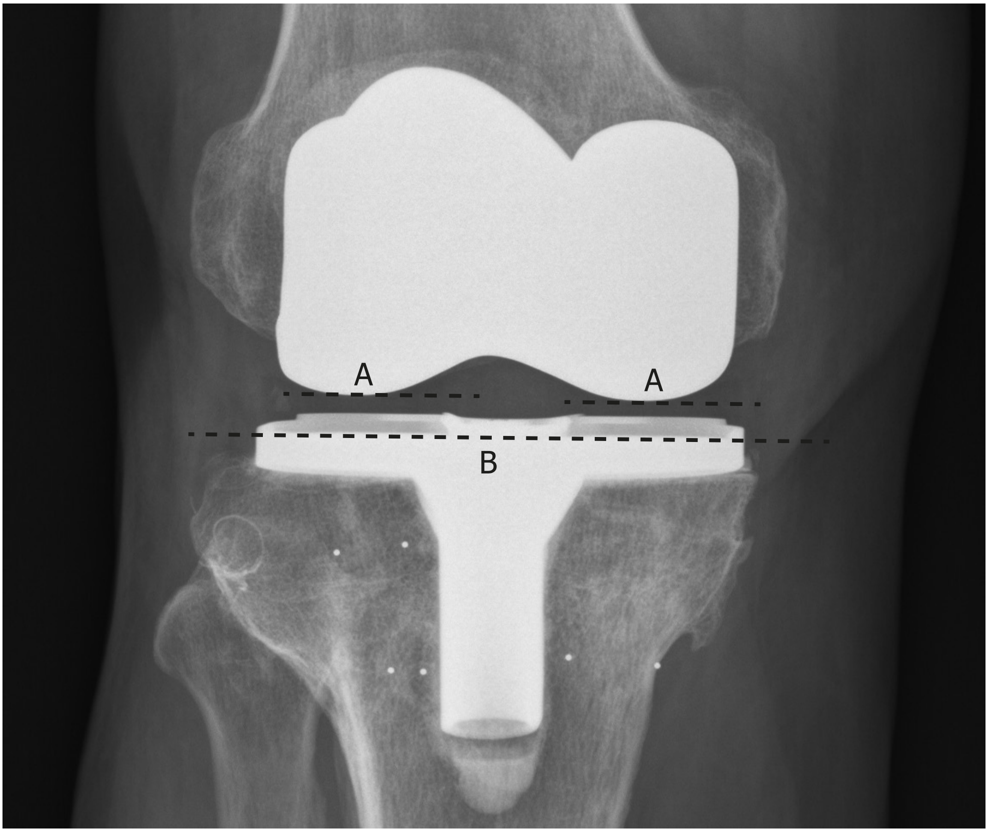


Figure 1-2. AP radiograph of a TKA with arrows indicating the lowest point of the femoral condyle (A) and the tibial tray (B), which are the reference points to assess the remaining insert thickness.

The conventional mJSW method is subject to parallax errors that occur when the metal tibial baseplate surface is not optimally aligned with the X-ray beam during sequential radiographic assessments. Moreover, some design features such as a metal rim require manual adjustments of the conventional mJSW measurement method, rendering the method sensitive for human errors. Measurement errors of up to 2 mm are not exceptional and multiple follow-up visits are required to obtain a reliable estimation of the wear rate [28, 29, 31]. These errors seriously limit the application of this measurement method for the purpose of reliably monitoring patients or evaluating implant designs.

1-6 Model-based wear measurement

Radiographic measurement accuracy and precision can be improved by applying model-based techniques. Such techniques incorporate prior information of three-dimensional (3D) object geometry and are applied to enhance clinical decision making or surgical accuracy by using computer-guided navigation. Model-based techniques are also applied in Roentgen Stereophotogrammetric Analysis, a method used to predict implant loosening after TKA or Total Hip Arthroplasty [32-34]. Accuracy and precision of 3D pose reconstruction have proven to be very high for these model-based techniques, therefore rendering them pre-eminently suitable for in vivo wear measurements [32].

The application of model-based techniques to the mJSW measurement has several advantages. Measurements applied in 3D are less susceptible to parallax errors than direct measurements in projection images of standard radiographs. Moreover, these techniques can improve signal-to-noise ratio because more image information is used when matching complete components compared to selecting a single image point or image edge. Last, using 3D models provides additional measurement possibilities, such as the location of the mJSW.

The model-based mJSW measurement can also be used to improve the mJSW measurement for the natural knee, where it is used to assess the progression of osteoarthritis[2, 35]. In this case, the 3D models should be capable of matching the patient-specific tibial and femoral shapes. This can be achieved by using deformable models that are able to match a variety of shapes and use smart fitting criteria to match only the desired shape. One of these models is the statistical shape model which uses a-priori knowledge from a training set to fit unseen shapesbased on their plausibility. This model type has proven successful in matching shapes of the natural knee based on the limited information available in projection images[36-38].

The use of model-based techniques also introduces new challenges. Apart from the need for accurate 3D (prosthesis) models, it requires a 3D reconstruction in which the spatial relationship between the projection image and the 3D model is established. To accomplish this, reliable information on the image acquisition process should be available, such as the original focus (camera) position with respect to the image, the image pixel size and the image magnification. In case this information is missing or unreliable, the precision of positioning the 3D models can drop quickly.

Although model-based techniques have been applied for mJSW measurements, the accuracy and precision of these techniques have not been validated or validation has been restricted to individual prostheses or other imaging modalities such as fluoroscopy and calibrated stereo imaging[30, 39-43].

1-7 Aim of this thesis

The aim of this work is to improve the accuracy and precision with which mJSW measurements can be conducted in medical imaging. Hereto, this thesis focusses on the development, validation and clinical application of model-based mJSW measurements for the natural knee and TKAs. For TKAs the measurement is applied for both stereo-images and standard radiographs.

1-8 Structure

Chapters 2 to 4 focus on the development and validation of polyethylene wear measurements for TKAs using calibrated stereo-images and 2D-3D matching of exact models. In Chapter 2 the accuracy of the mJSW method using model-based RSA is validated in a phantom study using different TKA designs. In Chapter 3 the differences in mJSW measurement between weight-bearing and non-weight-bearing stereo images are assessed. Alternative to the mJSW measurement, polyethylene wear can also be estimated using the wear volume. In Chapter 4 the precision of this volumetric wear measurement is analyzed by using both a phantom experiment and simulation studies.

Chapter 4 and 5 focus on the validation of the mJSW measurement in standard AP radiographs (i.e. mono images). In Chapter 4 a phantom study is used to perform this validation, which is equivalent to Chapter 1 for stereo-images. Ultimately, in vivo data are the most reliable basis to validate a measurement for clinical practice. In Chapter 5 a retrieval study is done to validate the measurement, in which the insert thickness measured in pre-revision images is compared to the actual insert thickness measured of the retrieved tibial inserts after revision.

Chapter 6 and 7 turn towards alternative model-based measurement techniques. In Chapter 6 the application of a volumetric wear measurement for knee prosthesis is considered and in Chapter 7 model-based techniques are applied to measure joint space narrowing in the natural knee. The accuracy and precision of the mJSW measurement of the knee using a Statistical Shape Model is compared to a conventional automated mJSW measurement.

Finally, Chapter 8 provides a general discussion and reflection on the improvement of the accuracy and precision with which mJSW measurements can be conducted in medical imaging. Also, directions for future work are described.

References

[1] Litwic, A., et al., *Epidemiology and burden of osteoarthritis.* British medical bulletin, 2013: p. lds038.

[2] Hellio Le Graverand, M.-P., et al., *Radiographic grading and measurement of joint space width in osteoarthritis.* Rheumatic diseases clinics of North America, 2009. **35**: p. 485-502.

[3] Poos, M.J.J.C. and A.M. Gommer. *Hoe vaak komt arthrose voor en hoeveel mensen sterven er aan?* 2014 23 juni 2014; Available from: http://www.nationaalkompas.nl/gezondheid-en-ziekte/ziekten-en-aandoeningen/bewegingsstelsel-en-bindweefsel/artrose.

[4] RIVM, *Kosten van Ziektenstudie*, 2013, RIVM.

[5] Leskinen, J., et al., *The incidence of knee arthrolasty for primary osteoarthritis*

grows rapidly among baby-boomers - a population-based study. Arthritis & Rheumatism, 2012. **64**: p. 423-428.

[6] Kurtz, S.M., et al., *Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030.* Clinical Orthopaedics and Related Research, 2009. **467**: p. 2606-12.

[7] Murphy, L., et al., *Lifetime risk of symptomatic knee osteoarthritis.* Arthritis Care & Research, 2008. **59**(9): p. 1207-1213.

[8] Wyatt, S.B., K.P. Winters, and P.M. Dubbert, *Overweight and obesity: prevalence, consequences, and causes of a growing public health problem.* The American Journal of the Medical Sciences, 2006. **331**(4): p. 166-74.

[9] *Swedish Knee Registry, Anual report 2011*. 28-08-2012 [cited 2012 August]; Available from: http://www.knee.nko.se/english/online/thePages/publication.php.

[10] Quinet, R.J. and E.G. Winters, *Total joint replacement of the hip and knee.* The Medical clinics of North America, 1992. **76**(5): p. 1235-51.

[11] Carr, A.J., et al., *Knee replacement.* The Lancet, 2012. **379**: p. 1331-1340.

[12] *Australian Orthopeadic Assocation National Joint Replacement Registry Anual report 2011*.

[13] *New Zealand National Joint Registry Annual Report 2009*.

[14] Robertsson, O., et al., *Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden.* Acta Orthopaedica Scandinavica, 2000. **71**: p. 262-7.

[15] Kurtz, S.M., et al., *Prevalence of primary and revision hip and knee arthroplasty in the United States from 1990 through 2002.* The Journal of Bone & Joint Surgery, 2005. **87**(7): p. 1487-1497.

[16] Naudie, D.D.R., et al., *Wear and osteolysis around total knee arthroplasty.* The Journal of the American Academy of Orthopaedic Surgeons, 2007. **15**: p. 53-64.

[17] Schmalzried, T.P. and J.J. Callaghan, *Current Concepts Review-Wear in Total Hip and Knee Replacements.* The Journal of Bone & Joint Surgery, 1999. **81**(1): p. 115-136.

[18] Lewis, G., *Polyethylene wear in total hip and knee arthroplasties.* Journal of Biomedical Materials Research, 1997. **38**: p. 55-75.

[19] Ingham, E. and J. Fisher, *The role of macrophages in osteolysis of total joint*

replacement. Biomaterials, 2005. **26**(11): p. 1271-86.

[20] Sundfeldt, M., et al., *Aseptic loosening, not only a question of wear: a review of different theories.* Acta Orthopaedica, 2006. **77**: p. 177-97.

[21] Mulhall, K.J., et al., *Current etiologies and modes of failure in total knee arthroplasty revision.* Clinical Orthopaedics and Related Research, 2006. **446**: p. 45-50.

[22] Sharkey, P.F., et al., *Why are total knee arthroplasties failing today?* Clinical Orthopaedics and Related Research, 2002. **404**: p. 7-13.

[23] Engh, C.A., et al., *Radiographically measured total knee wear is constant and predicts failure.* The Journal of Arthroplasty, 2013.

[24] Pijls, B.G., H.M. Van der Linden-Van der Zwaag, and R.G. Nelissen, *Polyethylene thickness is a risk factor for wear necessitating insert exchange.* International Orthopaedics, 2012. **36**(6): p. 1175-80.

[25] Dennis, D.A. and R.D. Komistek, *Mobile-bearing total knee arthroplasty: design factors in minimizing wear.* Clinical Orthopaedics and Related Research, 2006. **452**: p. 70-7.

[26] Lavernia, C.J., et al., *Activity level and wear in total knee arthroplasty: a study of autopsy retrieved specimens.* The Journal of Arthroplasty, 2001. **16**: p. 446-53.

[27] Miller, T.T., *Imaging of knee arthroplasty.* European Journal of Radiology, 2005. **54**: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, *Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses.* The Journal of Arthroplasty, 2003. **18**: p. 860-866.

[29] Collier, M.B., et al., *Radiographic assessment of the thickness lost from polyethylene tibial inserts that had been sterilized differently.* The Journal of Bone & Joint Surgery. American Volume, 2008. **90**: p. 1543-52.

[30] Sanzén, L., et al., *Radiographic wear assessment in a total knee prosthesis.* The Journal of Arthroplasty, 1996. **11**: p. 738-42.

[31] Mintz, A.D., C.A. Pilkington, and D.W. Howie, *A comparison of plain and fluoroscopically guided radiographs in the asessment of arthroplasty of the knee.* The Journal of Bone & Joint Surgery. American Volume, 1989. **71**: p. 1343-1347.

[32] Kaptein, B.L., et al., *A new model-based RSA method validated using CAD models and models from reversed engineering.* Journal of Biomechanics, 2003. **36**: p. 873-882.

[33] Kaptein, B.L., et al., *Clinical validation of model-based RSA for a total knee prosthesis.* Clinical Orthopaedics and Related Research, 2007. **464**: p. 205-9.

[34] Ryd, L., et al., *Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses.* The Journal of Bone & Joint Surgery. British Volume, 1995. **77**: p. 377-83.

[35] Vignon, E., et al., *Measurement of radiographic joint space width in the tibiofemoral compartment of the osteoarthritic knee: comparison of standing anteroposterior and Lyon schuss views.* Arthritis & Rheumatism, 2003. **48**: p. 378-84.

[36] Baka, N., et al., *2D-3D shape reconstruction of the distal femur from stereo X-Ray imaging using statistical shape models.* Medical Image Analysis, 2011. **15**(6): p. 840-850.

[37] Baka, N., et al., *Correspondence free 3D statistical shape model fitting to sparse x-ray projections.* SPIE, 2010: p. 76230D-76230D-9.

[38] Baldwin, M.a., et al., *Development of subject-specific and statistical shape models of the knee using an efficient segmentation and mesh-morphing approach.* Computer methods and programs in Biomedicine, 2010. **97**: p. 232-40.

[39] Burckhardt, K., et al., *Measuring orthopedic implant wear on standard radiographs with a precision in the 10 micrometer range.* Medical Image Analysis, 2006. **10**: p. 520-529.

[40] Duryea, J., et al., *Fully automated software to monitor wear in prosthetic knees using fluoroscopic images.* European Radiology, 2001. **11**: p. 2184-7.

[41] Fukuoka, Y., A. Hoshino, and A. Ishida, *A simple radiographic measurement method for polyethylene wear in total knee arthroplasty.* IEEE Transactions of Rehabilitation Engineering, 1999. **7**: p. 228-233.

[42] Kellett, C.F., et al., *In vivo measurement of total knee replacement wear.* The Knee, 2004. **11**: p. 183-187.

[43] Short, A., et al., *A novel method for in vivo knee prosthesis wear measurement.* Journal of Biomechanics, 2005. **38**: p. 315-322.

2

The robustness and accuracy of measuring the minimum joint space width of TKA based on model-based RSA

The robustness and accuracy of measuring the minimum joint space width of total knee arthroplasty based on model-based RSA

E.A. van IJsseldijk[a], E.R. Valstar[a,b], B.C. Stoel[c], R.G.H.H. Nelissen[a], J.H.C. Reiber[c], B.L. Kaptein[a]

[a] *Biomechanics and Imaging Group, Department of Orthopaedics, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*

[b] *Department of Biomechanical Engineering, Faculty of Mechanical, Maritime, and Materials Engineering, Delft University of Technology, Delft, The Netherlands*

[c] *Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands*

Published as "The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA", Journal of biomechanics, 2011, 44(15), pages 2724-2727

Abstract

Introduction

Accurate in vivo measurements methods of wear in total knee arthroplasty are required for a timely detection of excessive wear and to assess new implant designs. Component separation measurements based on model-based Roentgen stereophotogrammetric analysis (RSA), in which 3-dimensional reconstruction methods are used, have shown promising results, yet the robustness of these measurements is unknown. In this study, the accuracy and robustness of this measurement for clinical usage was assessed.

Method

The validation experiments were conducted in an RSA setup with a phantom setup of a knee in a vertical orientation. 72 RSA images were created using different variables for knee orientations, two prosthesis types (fixed-bearing Duracon knee and fixed-bearing Triathlon knee) and accuracies of the reconstruction models. The measurement error was determined for absolute and relative measurements and the effect of knee positioning and true separation distance was determined.

Results

The measurement method overestimated the separation distance with 0.1 mm on average. The precision of the method was 0.10 mm (2*SD) for the Duracon prosthesis and 0.20 mm for the Triathlon prosthesis. A slight difference in error was found between the measurements with 0° and 10° anterior tilt. (difference = 0.08 mm, $p = 0.04$).

Conclusion

The mJSW can be measured with an accuracy of 0.1 mm and precision of 0.2 mm based on model-based RSA, which is more than adequate for clinical applications. The measurement is robust in clinical settings. Although anterior tilt seems to influence the measurement, the size of this influence is low and clinically irrelevant.

2-1 Introduction

Total knee arthroplasty (TKA) is highly successful in relieving pain and restoring joint function, yet implant failure remains a problem. One of the main causes of failure is excessive polyethylene wear. Wear particles can induce osteolysis that may provoke complications such as aseptic loosening. It has been reported that wear and osteolysis are the primary indications for revision in more than 44% of all revisions performed more than two years after surgery [22]. Excessive wear is related to the design of a prosthesis [25]. Therefore, new prosthesis designs are assessed with knee simulator studies before market introduction. Unfortunately these studies are limited in incorporating important factors such as patient activity and the incidence of misalignment [16, 26]. As an alternative, model-based Roentgen stereophotogrammetric analysis (MBRSA) may be used to assess wear in a clinical setting. This imaging and analysis method achieves sub-millimeter precision in assessing migration of prostheses [44-47], which is used to predict prosthetic loosening [34]. Wear measurements can be obtained with MBRSA and high accuracies were already obtained [42, 43, 48]. However, validation of these wear measurements has been restricted to individual prostheses or measurement protocols. The method's robustness to variations in patient positioning has not been characterized.

The goal of this study is to determine the robustness of TKA wear measurements in MBRSA. The study uses an RSA setup and a knee phantom in which the separation distance between the tibial and femoral components is known exactly. The measurement method is applied for different variables such as prostheses type, actual separation distance, digital model accuracy and patient positioning. The robustness of the method is determined by assessing the measurement error as a function of these variables.

2-2 Materials and Methods

We now describe the phantom setup, the MBRSA analysis and the details of the separation measurements that were used in this study.

2-2-1 Phantom setup and acquisition of RSA images

A phantom setup was used with the total knee prostheses fixed into sawbones, to create more realistic images. RSA images of the phantom setup in standing position were acquired with a vertical RSA setup[32]. The setup consisted of a vertical rail on a base plate with two supports on which a tibial and a femoral sawbone could be fixed (Figure 2-3 Left). RSA images were obtained with two synchronized X-ray sources each aimed at a digital X-ray detector (Canon CXDI-series, 169dpi, 12BPP). The detectors were placed adjacently in a carbon calibration box (Medis Specials b.v., Leiden, Netherlands). The X-ray sources were positioned 1.5m from the detectors with a 40° angle between their respective beams. The phantom device was positioned as close to the detectors as possible (Figure 2-3 Right).

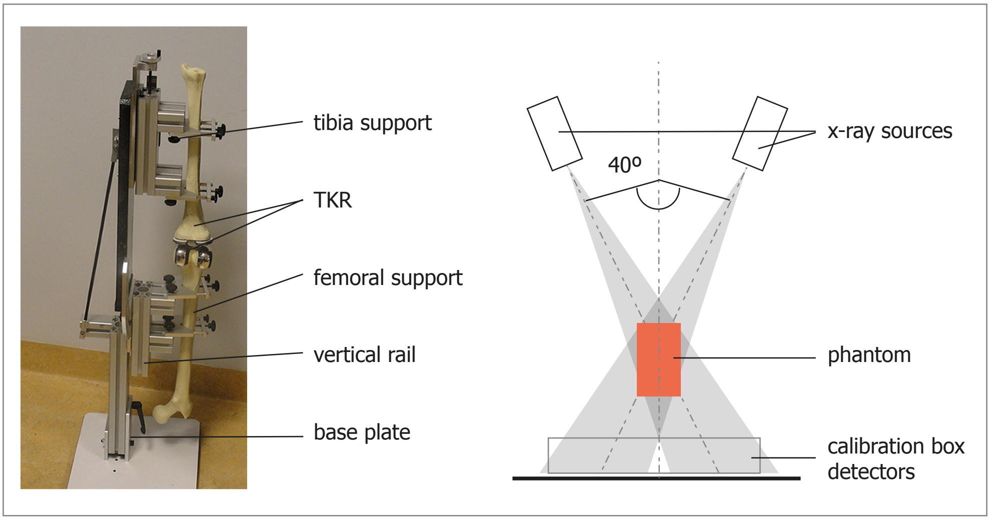


Figure 2-3. of the phantom-set-up. Right: Schematic top view of the RSA set-up

To validate the wear measurements, we analyzed the effect of different variables on the measurement error. In total 72 measurements were obtained using the variables in Table 2-1.

Prosthesis type

Two types of Stryker (Kalamazoo, USA) total knee prosthesis were used: the fixed-bearing Duracon knee (tibia size XL2, femur size XL) and the fixed-bearing Triathlon knee tibia (size 7, femur size 7).

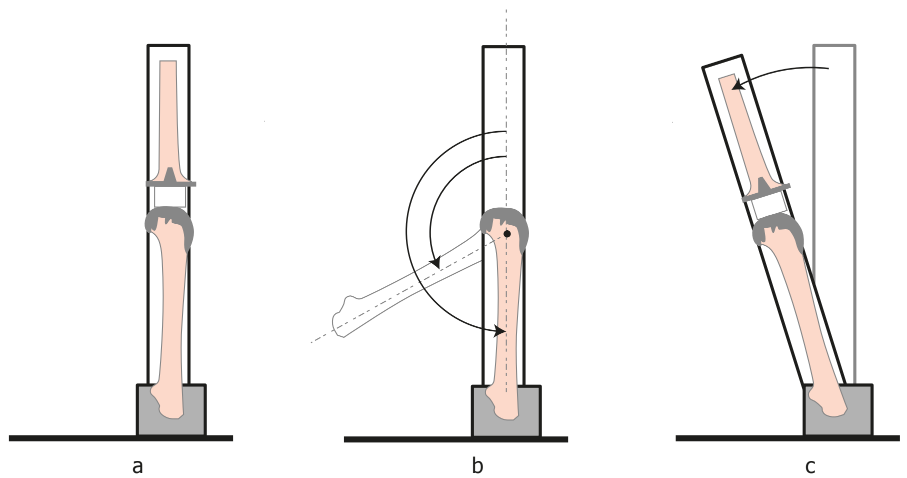


Figure 2-4. Phantom device settings: (a) in resting position, (b) with a flexion angle and (c) with a positive anterior tilt

Flexion angle, anterior tilt and rotation

To test for different flexion angles and the effect of patient positioning, the setup contained mechanisms to adapt the flexion angle of the knee, the anterior tilt and rotation of the leg with respect to the imaging system (Figure 2-3).

Component separation distance

The component separation distance was set using cylindrical, radiolucent plates (Plexiglass/PMMA), which had an accurate thickness (tolerance 0.05 mm). During the measurement a plate was placed in contact between the tibia plateau an d the medial femoral condyle of the total knee. By repeating the measurements with plates of 5 and 10 mm, we validated different component separation sizes.

2-2-2 Separation distance measurement

The separation distance measurement is based on 3D models of the tibial and femoral components. The first step of the measurement was creating a 3D reconstruction of the prosthesis component positions. An RSA analysis was done with MBRSA software

(Version 3.3, Medis Specials, Leiden, The Netherlands). The image contours of the components were selected semi-automatically. The user selected a region of interest in which the software program detected candidate edges (canny edge detection), which could be altered manually. Subsequently, the model poses were calculated by minimizing the difference between the image edges and the projected model silhouette. This is a standard procedure in MBRSA and the accuracy of the position and orientation estimation equaled 0.11 mm and 0.23°, respectively [33]. Next, the medial separation distance was calculated, which was defined as the shortest distance between the medial condyle of the femur and the tibial plane.

The RSA analyses were conducted with both computer aided design (CAD) models and models obtained by reverse engineering (RE), giving 144 measurement outcomes in total. The CAD models were provided by the prosthesis manufacturer. The RE models were created with a 3D laser scanner (Hyscan, Hymarc Tech, Ottawa, Canada) using the original components from this experiment. This scan had a tolerance of 0.020 mm.

Table 2-1. List of variables used in the robustness validation experiment

Order	Variable	Options	Procedure
1	Prosthesis type	1: Duracon 2: Triathlon	Place the sawbones with the prostheses components into the phantom setup
2	Flexion angle	1: 0° 2: 30° 3: 45°	Adapt the angle with the lever on the phantom setup (Figure 2-4b)
3	Separation distance	1: 10 mm 2: 5 mm	Fix the plate with the appropriate thickness between the tibial and femoral components
4	Anterior tilt	1: 0° 2: 10°	Adjust the angle with the phantom setup (Figure 2-4c)
5	Rotation	1: 0° 2: 10° 3: - 10°	Rotate the phantom device

2-2-3 Statistical analysis

The accuracy and precision of the measurement method were analyzed based on the measurement error, which is the difference between the measurement outcome and the separation distance set during the measurement. The means and standard deviations of the error were calculated for each subgroup of prosthesis type, model type and flexion angle. This was carried out to determine and compare systematic errors among these groups. Subsequently, tests were applied to determine whether mean errors were influenced by anterior tilt, actual separation distance and internal rotation (*t*-test/ANOVA). These tests were conducted with the data from RE models only, to avoid confounding due to model inaccuracies.

2-3 Results

Table 2-2 and Figure 2-5 show the average measurement error per group of prosthesis/model type and flexion angle. These groups consisted of 12 measurements combining all anterior tilt angles, rotation angles and separation distances.

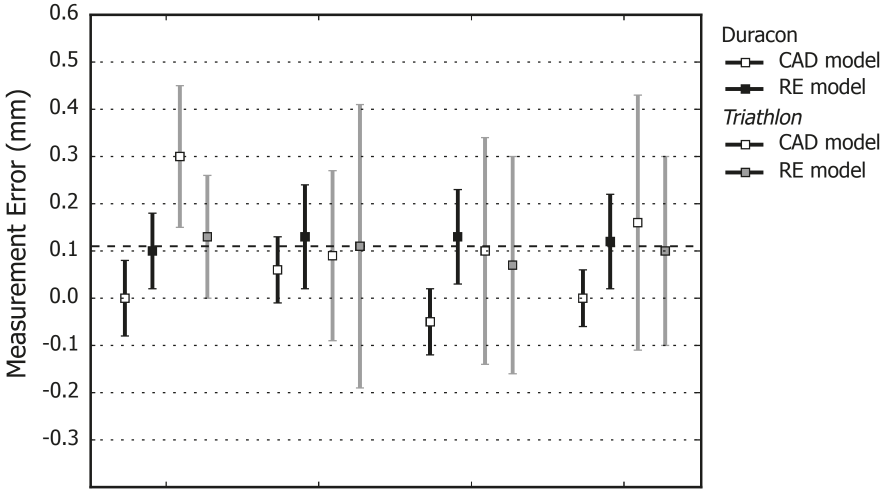


Figure 2-5. Measurement errors of the absolute wear measurement for different subgroups of prosthesis and flexion angles. Each group consists of 12 measurements. The dashed horizontal line shows the average measurement error (0.11 mm).

Table 2-2. Mean measurement errors in the robustness experiment, comparison between model types, prosthesis type and knee flexion angle. Each subgroup consistsof12 measurements. Subgroups with a significant error (p < 0.05, t-test) are printed in bold.

	Flexion angle (deg)		**0°** N = 12	**30°** N = 12	**45°** N = 12	**All** N = 36
Duracon	CAD	Mean + 2*SD	0.00 ± 0.08	**0.06 ± 0.07**	**-0.05 ± 0.07**	0.00 ± 0.06
		Median	- 0.012	0.046	- 0.058	0.005
		Total range	[-0.05 - 0.09]	[0.02 - 0.14]	[-0.09 - 0.01]	[-0.09 - 0.14]
	RE	Mean + 2*SD	**0.10 ± 0.08**	**0.13 ± 0.11**	**0.13 ± 0.10**	**0.12 ± 0.10**
		Median	0.09	0.12	0.12	0.11
		Total range	[0.05 – 0.19]	[0.06 – 0.22]	[0.08 – 0.21]	[0.05 – 0.22
Triathlon	CAD	Mean + 2*SD	**0.30 ± 0.15**	**0.90 ± 0.18**	**0.10 ± 0.24**	**0.16 ± 0.27**
		Median	0.31	0.06	0.11	0.18
		Total range	[0.19 – 0.40]	[-0.05 – 0.23]	[-0.12 – 0.29]	[-0.12 – 0.40]
	RE	Mean + 2*SD	**0.13 ± 0.22**	**0.11 ± 0.20**	**0.07 ± 0.23**	**0.10 ± 0.20**
		Median	0.15	0.15	0.05	0.09
		Total range	[-0.02 – 0.26]	[-0.06 – 0.20]	[-0.04 – 0.26]	[-0.06 – 0.26]

The results indicated that a systematic overestimation error of 0.1 mm was present in general (one sample *t*-test, $p < 0.05$) and in 11 out of 12 subgroups. As can be seen in Figure 2-5, the error of measurement with CAD models varied significantly over the flexion angles for both prosthesis types (ANOVA, $p < 0.001$). Measurements with RE models did not show this variation.

As shown in Table 2-2, the measurements were more precise with the Duracon prosthesis than with the Triathlon prosthesis (0.2 and 0.1 mm, respectively for the RE models). Levene's test was applied on 6 equivalent subgroups (flexion*model type) and the outcome was significant ($p < 0.05$) for all but the Duracon 0° flexion case.

The mean errors between the groups of anterior tilt, knee rotation and real separation distance are displayed in Table 2-3. Only for anterior a significant difference in error was found (d=0.08, *t*-test, $p < 0.05$).

Table 2-3. Average measurement errors and standard deviations (SD) for different values for anterior tilt, knee rotation and real component separation. The last column shows the difference test used and result for significance.

		N	Average (mm)	SD (mm)	Difference test
Anterior tilt	0°	36	0.07	0.07	*t*-test
	10°	36	0.15	0.06	p = 0.04
Separation distance	5 mm	36	0.09	0.09	*t*-test
	10 mm	36	0.13	0.07	p = 0.06
Internal rotation	-10°	24	0.09	0.08	ANOVA
	0°	24	0.11	0.07	p = 0.11
	10°	24	0.14	0.08	

2-4 Discussion

We studied the accuracy and precision of a component separation measurement in MBRSA for TKA. The study was performed with a phantom setup, in which the measurement was repeated for various knee positions, separation distances and prosthesis types.

We found that the measurement had a small overestimation of 0.1 mm. For the CAD models, this seems to depend on the flexion angle, whereas the results for the RE models were more homogeneous. In addition, anterior tilt may influence the measurement, as a statistically significant effect size of 0.07 mm was observed over a tilt range of 10°. However, this effect is small and should pose little concern when patients are positioned carefully.

Other similar wear validation studies using perspex/acrylic plates also reported overestimations [43, 49]. We noticed that in many measurements the image contours of the prostheses were systematically smaller/more contracted than the contours based on the model projections. This difference may lead to a systematic error in the pose estimation, putting the models further apart and thus increasing the measured separation distance. This error is neutralized in relative measurements over time such as migration, which are usually performed in MBRSA studies.

The precision of the measurement seems related to the prosthesis type. With RE models, the precision of the Triathlon and Duracon prosthesis were 0.2 and 0.1 mm, respectively. Possibly, the Duracon prosthesis has a more salient geometry, giving a higher precision in the pose estimation.

An important question is how these results influence a wear measurement, which is the difference between two subsequent separation distance measurements. Assuming these measurements are independent, the overestimations will cancel out and a precision is expected of $\sqrt{2} * 0.2 \approx 0.3$ mm. This shows the measurement is suitable for clinical research studies, as sub-millimeter difference can be detected with small patient groups.

A limitation of this study is the lack of experiments with in vivo data, in which soft tissue attenuation can deteriorate contours detection. Still, similar results are expected as attenuation is usually limited in knee X-rays. Besides, MBRSA analysis is robust even if only 10% of the contour information is used [50].

Some general limitations still exist for TKA wear measurements based on the separation distance. Wear is localized and liners can have a congruent geometry [25, 51]. Therefore, the outcome of the measurement depends on the contact location of the femur, which decreases the reproducibility of the measurement *in vivo*. In addition, the measurement cannot distinguish between wear and creep. Creep stabilizes in the first years after surgery [52, 53], after which period the wear measurement becomes reliable.

In conclusion, our data shows that the joint separation measurements based on model-based RSA are accurate enough for wear studies of total knee prostheses. Further research is needed for the usability in clinical practice. The use of RE models is recommended, as the measurement is more robust compared to CAD.

2-5 Acknowledgements

The authors wish to thank the Radiology Department the Leiden University Medical Center for their assistance with this study. The authors with to thank the Instrumental

Department at the Leiden University Medical Center for their support in developing the phantom setup. Funding for this study was provided by the Dutch Arthritis Foundation (no. 08-1-303) and Stryker.

References

[16] Naudie, D.D.R., et al., *Wear and osteolysis around total knee arthroplasty.* The Journal of the American Academy of Orthopaedic Surgeons, 2007. **15**: p. 53-64.

[22] Sharkey, P.F., et al., *Why are total knee arthroplasties failing today?* Clinical Orthopaedics and Related Research, 2002. **404**: p. 7-13.

[25] Dennis, D.A. and R.D. Komistek, *Mobile-bearing total knee arthroplasty: design factors in minimizing wear.* Clinical Orthopaedics and Related Research, 2006. **452**: p. 70-7.

[26] Lavernia, C.J., et al., *Activity level and wear in total knee arthroplasty: a study of autopsy retrieved specimens.* The Journal of Arthroplasty, 2001. **16**: p. 446-53.

[32] Kaptein, B.L., et al., *A new model-based RSA method validated using CAD models and models from reversed engineering.* Journal of Biomechanics, 2003. **36**: p. 873-882.

[33] Kaptein, B.L., et al., *Clinical validation of model-based RSA for a total knee prosthesis.* Clinical Orthopaedics and Related Research, 2007. **464**: p. 205-9.

[34] Ryd, L., et al., *Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses.* The Journal of Bone & Joint Surgery. British Volume, 1995. **77**: p. 377-83.

[42] Kellett, C.F., et al., *In vivo measurement of total knee replacement wear.* The Knee, 2004. **11**: p. 183-187.

[43] Short, A., et al., *A novel method for in vivo knee prosthesis wear measurement.* Journal of Biomechanics, 2005. **38**: p. 315-322.

[44] Garling, E.H., E.R. Valstar, and R.G.H.H. Nelissen, *Comparison of micromotion in mobile bearing and posterior stabilized total knee prostheses: a randomized RSA study of 40 knees followed for 2 years.* Acta Orthopaedica, 2005. **76**: p. 353-61.

[45] Nelissen, R.G.H.H., *Fundamental aspects of the clinication evalution of total*

knee prostheses, 1995.

[46] Nilsson, K.G. and J. Kärrholm, *RSA in the assessment of aseptic loosening.* The Journal of Bone & Joint Surgery. British Volume, 1996. **78**: p. 1-3.

[47] Soballe, K., et al., *Migration of hydroxyapatite coated femoral prostheses. A Roentgen Stereophotogrammetric study.* The Journal of Bone & Joint Surgery. British Volume, 1993. **75**: p. 681.

[48] Gill, H.S., et al., *In vivo measurement of volumetric wear of a total knee replacement.* The Knee, 2006. **13**: p. 312-7.

[49] Simpson, D.J., et al., *Development of a model-based RSA systeem to measure polyethylene wear in unicompartmental arthroplasty.* Clinical Biomechanics, 2010. **?**: p.?

[50] Kaptein, B.L., et al., *Evaluation of three pose estimation algorithms for model-based Roentgen stereophotogrammetric analysis*, in *The Journal of Engineering in Medicine, part H*2004. p. 231-8.

[51] Bartel, D.L., V.L. Bicknell, and T.M. Wright, *The effect of conformity, thickness, and material on stresses in ultra-high molecular weight components for total joint replacement.* The Journal of Bone & Joint Surgery. American Volume, 1986. **68**: p. 1041-1051.

[52] Deng, M., et al., *Study of creep behavior of ultra-high-molecular-weight polyethylene systems.* Journal of Biomedical Materials Research, 1998. **40**: p. 214-223.

[53] Lee, K.-Y. and D. Pienkowski, *Compressive creep characteristics of extruded ultrahigh-molecular-weight polyethylene.* Journal of Biomedical Materials Research, 1998. **39**: p. 261-265.

3

A model-based approach to measure the minimum joint space width of TKA in standard radiographs

A model-based approach to measure the minimum joint space width of total knee arthroplasty in standard radiographs

E.A. van IJsseldijk[a], E.R. Valstar[a,b], B.C. Stoel[c], R.G.H.H. Nelissen[a], B.L. Kaptein[a]

[a] *Biomechanics and Imaging Group, Department of Orthopaedics, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*

[b] *Department of Biomechanical Engineering, Faculty of Mechanical, Maritime, and Materials Engineering, Delft University of Technology, Delft, The Netherlands*

[c] *Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands*

Published as "A model-based approach to measure the minimjm joint space width of total knee replacements in standard radiographs", Journal of biomechanics, 2012, 45(12), pages 2171-2175

Abstract

Introduction

Excessive wear is in total knee arthroplasty is detected by measuring the minimum joint space width (mJSW) in anterioposterior radiographs. The accuracy of conventional measurement methods is limited and can be improved using model-based techniques. In this study, the model-based wear measurement (MBWM) is introduced. Its accuracy and reproducibility are assessed and compared to the conventional measurement.

Method

40 anterioposterior radiographs were obtained of a knee prosthesis using a phantom set-up. Both measurement methods were applied and the accuracy and precision were compared. The reproducibility was calculated with an inter- and intra-observer experiment. Three observers measured the mJSW in 30 clinical radiographs with both the conventional measurement and the MBWM and repeated this after 6 weeks. The experiments were conducted with a NexGen mobile bearing and fixed bearing prostheses.

Results

In the phantom experiment, the accuracy (mean of the absolute error) was significantly higher (t-test, $p < 0.01$) for the MBWM as for the conventional measurement (0.15 mm versus 0.43 mm, 0.14 mm versus 0.35 mm for the mobile and fixed bearing respectively). The standard deviation of the measurements is smallest for the MBWM measurement for both prosthesis types (0.16 mm versus 0.47 mm, Levene's test, $p < 0.01$). In the reproducibility experiment, both the intra- and inter-observer agreements was higher for the MBWM than for the conventional method.

Conclusions

The results show that the MBWM is superior to the conventional measurement in both accuracy and reproducibility. Although the use of a phantom experiment poses some limitations in conveying the findings to clinical practice, this improved mJSW measurement can lead to better wear detection for surgery decisions and research purposes.

3-1 Introduction

Excessive polyethylene wear is an important cause of implant failure in total knee arthroplasty (TKA)[16, 22]. As the incidence of total knee arthroplasty is increasing, the impact of wear problems is expected to increase as well[54].

In current clinical practice, polyethylene wear is determined in vivo using the minimum joint space width (mJSW), which is assessed in standard radiographs. This diagnostic tool is used to evaluate new prosthesis designs and for decision support for surgical procedures such as isolated polyethylene exchanges[55-57].
The mJSW is obtained in anterioposterior (AP) or mediolateral radiographs[27, 28]. However, the accuracy is limited and measurement errors higher than 1 mm are not exceptional[28].

The measurement accuracy and precision can be improved by model-based techniques. In our previous work, we described and validated a wear measurement method for model-based roentgen stereophotogrammetric analysis (MBRSA), in which the tibia-femoral distance is obtained based on 3D surface models of the components[58], using 3D vision techniques[59].

This approach can be used for standard radiographs as well. The accuracy of the measurement will be lower than the accuracy found in MBRSA, as accurate calibration is not possible and model matching is done with a single X-ray source only. Nonetheless, we hypothesize that the generally applicable, model-based approach will be more accurate and reproducible compared to conventional methods, as more image information is exploited and less dependency is expected to joint space narrowing caused by anterioposterior tilt of the tibial baseplate.

The goal of this study is to evaluate the accuracy and reproducibility of the model-based wear measurements in AP radiographs (MBWM) and compare the results to the conventional measurement method. To determine the accuracy, a phantom set-up was constructed with a known mJSW, in which the measurements were conducted for different positions of the phantom, insert sizes and prosthesis types. The reproducibility was determined using inter- and intra-observer studies with clinical data.

3-2 Materials and Methods

3-2-1 Measurement Methods

We now describe the metal-to-middle (conventional) and model-based measurements. Both methods determine the mJSW as the shortest distance between the tibial tray and the femoral condyles. The first method uses the visible distance in the image itself, whereas the second uses a semi-automatic measurement based on 3-D models that are matched with the image.

Metal-to-middle measurement

The metal-to-middle measurement is the standard method in obtaining the mJSW in the image [27, 28]. A reference line is drawn through the tibial tray at its largest medialateral width. Then, the shortest perpendicular distances are estimated between this reference and the femoral condyles (Figure 3-6 left).

In our experiments, the metal-to-middle method was conducted using a computer software (Digimizer® version 4.0.0.0, MedCalc Software, Mariakerke, Belgium). The image magnification was corrected using the ratio of the known width of the tibial baseplate to the width in the radiograph.

Model-based wear measurement

In the model-based wear measurement (MBWM) method, 2-D/3-D registration is used to match 3-D surface models of the tibial and femur components with the AP radiograph. Then, the minimal medial and lateral distances are automatically measured based on the models (Figure 3-6 right).

The 2-D/3-D registration was conducted in model-based RSA (Version 3.3, Medis Specials, Leiden, The Netherlands). The origin of the laboratory frame was located in the center of the image detector. The *x* and *y* coordinates thereby describe the image plane, whereas the *z* coordinate is the direction perpendicular to the image. It was assumed that the position of the X-ray source was located on the *z*-axis (e.g. perpendicular to the image center). The DICOM information was used to set the distance between the X-ray source and the detector and the physical pixel spacing of the detector.

The image contours of the tibial and femoral components were selected semi-automatically with canny edge detector [60]. The position and orientation of the models were calculated by minimizing the difference between the image contours and the projected model silhouettes.

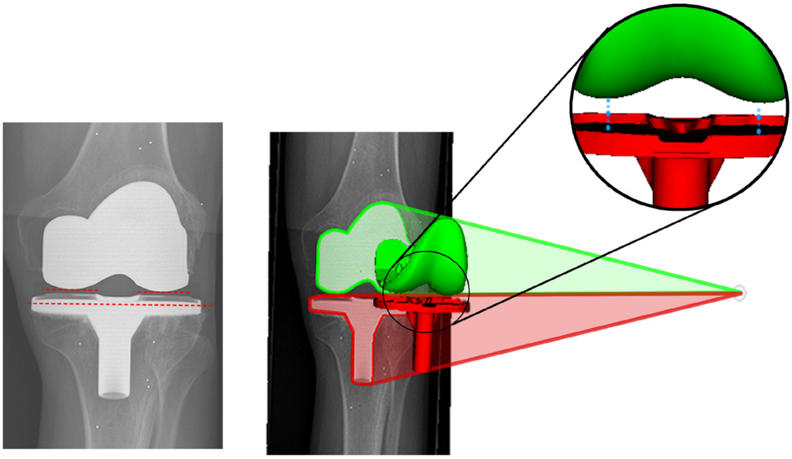


Figure 3-6. Comparison of the measurement methods: (left) the metal-to-middle method, in which the mJSW is obtained in teh radiograph, and (right) the MBWM in which the mJSW is obtained semi-automatically based on 3D models

3-2-2 Experiments

Phantom experiment

The phantom setup consisted of the tibial and femoral components of the knee prostheses, which were inserted into sawbones (Figure 3-7). The setup was placed in an X-ray imaging system (CXDI-series, 169dpi, 12BPP, Canon, New York, USA), according to the anterior-posterior (AP) protocol in standing position as used in our hospital. The X-ray source was positioned 1.2 meters from the detectors and the phantom was positioned approximately 20 cm from the detector.

The actual mJSW was set using radiolucent plates (Plexiglass/PMMA), which had an accurately defined thickness (tolerance 0.05 mm). Four different sizes (5, 8, 10

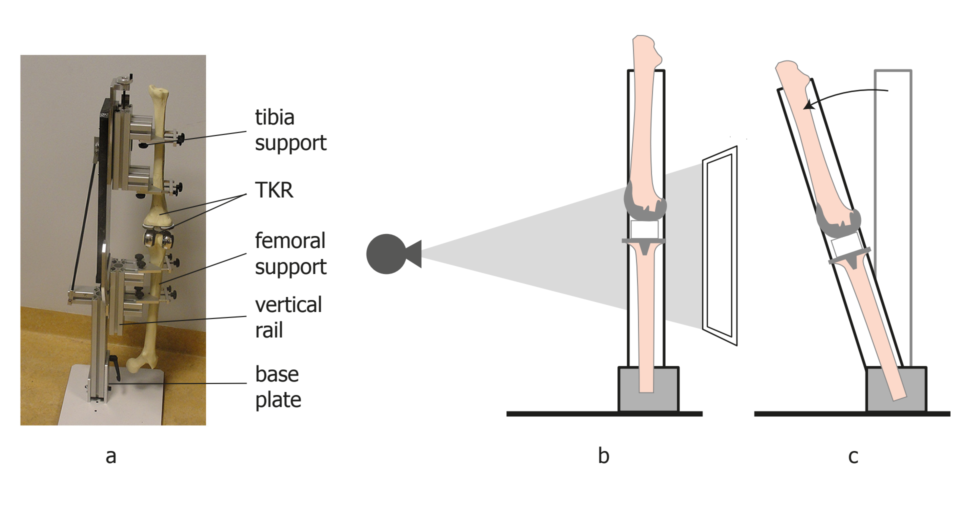


Figure 3-7. a) image of the phantom. (b) illustration of the phantom in a neutral position. (c) the phantom with anterior tilt.

and 12 mm) were used and the appropriate plate was placed between the tibial tray and the medial femoral condyle during the acquisition. As contact was possible only for the medial condyle, the lateral mJSW was not measured in the experiment.

For each plate size 10 images were acquired. Among these images, the position (range -10 to 10 cm) and orientation (range -10° to 10°) of the phantom with respect to the image were varied. In addition, the setup was placed in different anterior tilt angles (range 0° to 10°), as illustrated in Figure 3-7.

We repeated the acquisitions for both the fixed bearing and mobile bearing NexGen (Zimmer, Warsaw, IN, USA) total knee prostheses to cover different geometric designs. The size of the fixed and mobile bearings were (5-F) and (4-D), respectively (tibia-femur). Computer aided design (CAD) models were available for all components, except for the tibia component of the fixed prosthesis. For this component a reversed engineered model was created with a 3-D laser scan (*Hyscan*, Hymarc Tech, Ottawa, Canada), which had a tolerance of 0.020 mm.

In total, 80 images were acquired (10 images x 4 plate sizes x 2 prosthesis types). For all images, the mJSW was calculated with the MBWM and conventional measurement method, which was obtained by a clinician. Subsequently, the errors were calculated as the difference with the actual mJSW defined by the plate thickness.

Statistical analysis

We calculated the error mean, absolute error mean, standard deviation of the error and error range per prosthesis type and measurement method. The sizes of the errors are tested for statistical significance with unpaired *t*-tests. Levene's tests are used to test for differences in variance between the errors of the measurement methods. Finally, the dependency between the error and actual size was determined using a regression analysis (Pearson's rho).

Clinical experiment

In this experiment, a comparison was made between the inter- and intra-observer variability of the conventional wear measurement and MBWM. Clinical data was used as no ground truth value is required to obtain this measure.

For both the mobile and the fixed bearing prosthesis, 15 bearing AP radiographs were retrieved from the hospital database, in a random order. Both bilateral and unilateral images were included.

Three observers were included in the experiment: a clinician, a researcher and a senior researcher. They were asked to measure the medial and lateral insert thickness in the radiographs using both the conventional and model-based measurement methods. Observers could practice until they felt comfortable with the methods, preventing learning curve effects. To obtain the intra-observer variability, the observers repeated the measurements after a period of at least 6 weeks. In this series, the average measurement duration of the model-based method was also recorded.

Statistical analysis

The inter- and intra-observer variability of each measurement method was analyzed with the interclass correlation coefficient (ICC, two-way random, absolute agreement). The difference in spread between the measurements methods was tested with Pitman's test for correlated measures [61]. Bland-Altman plots were created to detect possible trends in the data [62].

3-3 Results

3-3-1 Phantom Experiment (accuracy)

The measurement errors (i.e. the difference between the measured thickness and the actual thickness of the plate) for each measurement method and prosthesis type are shown in box plots (Figure 3-8). The statistical characteristics of the errors are shown in Table 3-4, split for the measurement method and prosthesis type.

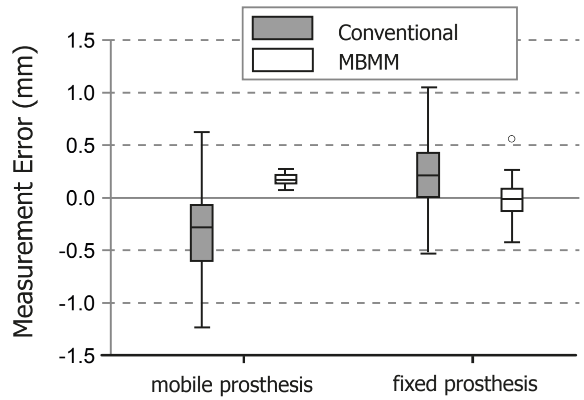


Figure 3-8. Boxplots showing the measurement errors of the different methods and prostheses types

Table 3-4. Statistics over the measurement errors in the phantom experiment.

	Mobile (N=40)		**fixed (N=40)**	
	conv	MBWM	conv	MBWM
mean (mm)	-0.36*	0.15*	0.20*	-0.03
standard deviation (mm)	0.40	0.06†	0.40	0.19†
absolute error mean (mm)	0.43	0.15**	0.35	0.14**
range (mm)	1.90	0.21	1.61	1.01

conv = conventional measurement; MBWM = model-based wear measurement

* statistically significant difference from 0 mm (*t*-test, $p < 0.05$)

† statistically significant difference in variance compared to the conventional measurement (Levene's test, $p < 0.01$)

** statistically significant difference in mean compared to the conventional measurement (*t*-test, $p < 0.01$)

For both prosthesis types, the standard deviation of the measurements is significantly smaller for the model-based measurement than for the conventional measurement (Levene's test, $p < 0.01$). Also, the model-based measurement had a significantly smaller standard deviation for the mobile prosthesis than for fixed prosthesis (Levene's test, $p < 0.01$). This is probably due to the implant geometry. The fixed prosthesis type contains thin structures such as the metal rim. These structures produce less pronounced image edges, increasing the localization error.

The average of the measurement error indicates whether a systematic bias (information bias) is present. The data from Table 3-4 shows that only the model-based measurement is unbiased for the fixed prosthesis. For the mobile prosthesis, the model-based measurement shows the smallest bias of the two methods. For both prosthesis types, the absolute error is lower for the MBWM than for the conventional measurement ($p < 0.01$).

No statistically significant correlation was found between the measurement error of the model-based measurement and the true distance for both prosthesis types (Pearson's rho, $p < 0.05$).

3-3-2 Clinical experiment (reproducibility)

The ICC values found in the inter- and intra-observer study were higher for the model-based method than for the conventional method for any observer and prosthesis type (Table 3-5), indicating a better reproducibility for the first method. On average, the reproducibility was higher for the fixed bearing than for the mobile bearing.

The average measurement durations of the observers in the second measurement series were 1:37, 2:24 and 2:37 (min:sec).

The standard deviations over all measurements with the conventional method and model-based method are 0.37 mm and 0.15 mm respectively. Pitman's test flagged the difference in spread significant ($p < 0.05$) for two out of three observers.

Table 3-5. Results of the inter and intra-observer variability of the measurement methods for the conventional (conv) and model-based method (mb) in terms of the interclass correlation coefficient.

	Intra-observer variability						**Interobserver variability**	
	Observer 1		**Observer 2**		**Observer 3**			
ICC values	conv	MBWM	conv	MBWM	conv	MBWM	conv	MBWM
mobile prosthesis	.945	.963	.926	.983	.822	.947	.863	.966
fixed prosthesis	.963	.986	.973	.982	.970	.992	.919	.982

The Bland-Altman plot gives the agreement between the measurements, by plotting the difference between the measurement methods (MBWM - conventional) against the mean value (Figure 3-9). Only the first measurement series was used and the mean value over the three observers was used, reducing the data to 60 points. This reduction keeps the plot legible and – more importantly - prevents oversampling, because the observer data contains a high dependency.

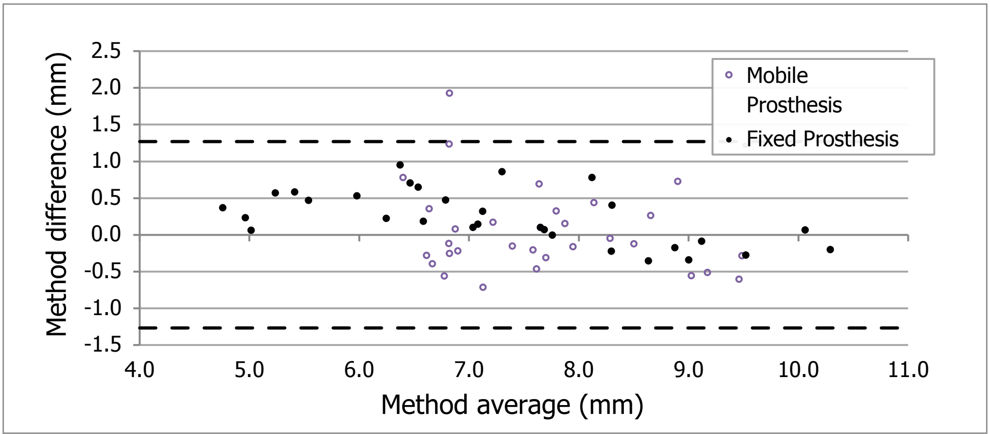


Figure 3-9. Blant-Altman plot showing the measurement agreement between the two methods.

The limit of agreement between the measurements is 1.27 mm (2 x SD), which is indicated with the broken lines. For the fixed prosthesis, the a statistically significant difference of 0.23 mm was found, i.e. the MBWM gives a larger value than the conventional measurement on average (*t*-test, p = 0.008).

3-4 Discussion

We developed a model based wear measurement (MBWM) that is superior in obtaining the mJSW in comparison with the method currently used in clinical practice. The main advantage is the improvement in precision and reproducibility that was obtained. A higher precision was found in the phantom experiment and a higher inter- and intra-observer reproducibility was found with clinical data. As these experiments were conducted with both a fixed and a mobile prosthesis design, we expect that these findings are generally applicable to other prosthesis designs as well.

Furthermore, we found a lower bias for the MBWM than for the conventional measurement in the phantom experiment. However, the high variability of the conventional measurements makes generalization of this result difficult. Furthermore, bias is of lesser importance than precision, because bias can nullify in relative measurements such as wear-rate measurements.

The average measurement duration of the model-based measurement was approximately 2 minutes, which is adequate for clinical use. We expect that this duration can be decreased by further automation of the contour detection, as the implant shadows are clearly distinguishable in AP radiographs.

The phantom experiment had several limitations. It did not include the soft tissue attenuation that is present in real clinical images. Still, the attenuation is usually limited and the pose estimator remains robust when only 5% of the complete contour is used [63]. Another limitation was that the geometric design of the fixed bearing tibia component in the phantom was different from the design in clinical images, due to this availability. This could have influenced the conventional measurement, because of differences in the metal rim surrounding the tibia plateau. This could explain why the mean difference between the measurements differs for this prosthesis, when the phantom experiment and clinical experiment are compared (-0.23 mm vs. 0.12 mm).

In clinical practice, the outcomes of the mJSW measurements (both conventional and model-based) depend on the articulating points of the femoral component at the moment of X-ray acquisition. Unfortunately, it was not possible to include this effect in the phantom experiment. Although we expect that congruent liners limit the variability of the articulating point, we hope to eliminate any uncertainty with a retrieval study, in which clinical radiographs are compared directly to retrieved liners. A general limitation of radiograph-based distance measurement is that creep and true wear cannot be distinguished. Instead, it is assumed that creep stabilizes within two months, whereas wear is expected to be a constant process over time [52].

Several other studies describe alternatives to the radiographic wear measurement. Some studies use fluoroscopy to improve the reproducibility, as the alignment between the tibial tray and the radiographic beam can be optimized before the measurement [40, 64]. A standard deviation of 0.15 mm was found in this measurement, which is similar to the finding in our work, yet fluoroscopy generally comes with a higher radiation dose for the patient and requires a longer imaging time. In other researches, a similar model-based wear measurement for calibrated single-source radiographs is described [41, 65]. Although the validation data is limited, this indeed seems to give better results (SD = 0.1 mm). However, this method imposes the presence of a calibration object. We think that the applicability to standard radiographs is a considerable advantage of the method we are using.

Based on these results, we conclude that the model-based method is a reliable tool to evaluate the insert thickness in standard radiographs. It can therefore aid in a better timing of insert exchanges, with the aim of decreasing the number of complications. Moreover, the accuracy of the method combined with the advantage that any standard radiograph can be used renders the method interesting for wear studies to compare prosthesis types.

References

[16] Naudie, D.D.R., et al., *Wear and osteolysis around total knee arthroplasty.* The Journal of the American Academy of Orthopaedic Surgeons, 2007. **15**: p. 53-64.

[22] Sharkey, P.F., et al., *Why are total knee arthroplasties failing today?* Clinical Orthopaedics and Related Research, 2002. **404**: p. 7-13.

[27] Miller, T.T., *Imaging of knee arthroplasty.* European Journal of Radiology,

2005. **54**: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, *Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses.* The Journal of Arthroplasty, 2003. **18**: p. 860-866.

[40] Duryea, J., et al., *Fully automated software to monitor wear in prosthetic knees using fluoroscopic images.* European Radiology, 2001. **11**: p. 2184-7.

[41] Fukuoka, Y., A. Hoshino, and A. Ishida, *A simple radiographic measurement method for polyethylene wear in total knee arthroplasty.* IEEE Transactions of Rehabilitation Engineering, 1999. **7**: p. 228-233.

[52] Deng, M., et al., *Study of creep behavior of ultra-high-molecular-weight polyethylene systems.* Journal of Biomedical Materials Research, 1998. **40**: p. 214-223.

[54] Robertsson, O., et al., *Past Incidence and future demands for knee arthroplasty in Sweden.* Acta Orthopaedica, 2000. **71**: p. 376-380.

[55] Bert, J.M., et al., *The incidence of modular tibial polyethylene insert exchange in total knee arthroplasty when polyethylene failure occurs.* The Journal of Arthroplasty, 1998. **13**: p. 609-14.

[56] Babis, G.C., R.T. Trousdale, and B.F. Morrey, *The effectiveness of isolated tibial insert exchange in revision total knee arthroplasty.* The Journal of Bone & Joint Surgery. American Volume, 2002. **84-A**: p. 64-8.

[57] Engh, G.a., L.M. Koralewicz, and T.R. Pereles, *Clinical results of modular polyethylene insert exchange with retention of total knee arthroplasty components.* The Journal of Bone & Joint Surgery. American Volume, 2000. **82**: p. 516-23.

[58] IJsseldijk, E.A.v., et al., *The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA.* Journal of Biomechanics, 2011. **44**(15): p. 2724-2727.

[59] Wunsch, P. and G. Hirzinger, *Registration of CAD-models to images by iterative inverse perspective matching*. Proceedings of 13th International Conference on Pattern Recognition1996: IEEE. 78-83 vol.1.

[60] Canny, J., *A Computational Approach To Edge Detection.* Pattern Analysis and Machine Intelligence, IEEE transactions on, 1986. **PAMI-8**(6): p. 679-698.

[61] PITMAN, E.J.G., *A note on normal correlation.* Biometrika, 1939. **31**: p. 9-12.

[62] Bland, J.M. and D.G. Altman, *STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT.* The Lancet, 1986. **327**(8476): p. 307-310.

[63] Kaptein, B.L., et al., *Evaluation of three pose estimation algorithms for model-based Roentgen stereophotogrammetric analysis.* J. Engineering in Medicine, part H, 2004: p. 231-8.

[64] Hide, I.G., et al., *A radiological technique for the assessment of wear in prosthetic knee replacements.* Skeletal Radiology, 2000. **29**: p. 583-586.

[65] Hoshino, A., A. Ishida, and Y. Fukuoka, *Accurate in vivo measurement of polyethylene wear in total knee arthroplasty.* The Journal of Arthroplasty, 2002. **17**: p. 490-496.

4

Measuring the minimum joint space width in total knee arthroplasty by RSA

Measuring the minimum joint space width in total knee arthroplasty by RSA

E.A. van IJsseldijk[a], E.R. Valstar[a,b], B.C. Stoel[c], R. de Ridder[d], R.G.H.H. Nelissen[a], B.L. Kaptein[a]

[a] *Biomechanics and Imaging Group, Department of Orthopaedics, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*

[b] *Department of Biomechanical Engineering, Faculty of Mechanical, Maritime, and Materials Engineering, Delft University of Technology, Delft, The Netherlands*

[c] *Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands*

[d] *Department of Orthopaedics, Langeland Ziekenhuis, Zoetermeer, The Netherlands*

Published as "Measuring Polyethylene wear in Total Knee Arthroplasty by RSA: Differences between Weight-Bearing and Non-Weight-Bearing Positioning", Journal of Orthopaedic Research, 2014, 32(4), pages 613-617

Abstract

Introduction

Measuring the minimum joint space width (mJSW) in total knee arthroplasty (TKA) in Roentgen Stereophotogrammetric Analysis (RSA) studies provides valuable information on polyethylene wear, a leading cause for TKA failure. Most existing RSA studies use non-weight-bearing (NWB) patient positioning. The latter may compromise mJSW measurements due to knee laxity with subsequent non-contact between the TKA components. We investigated the difference in mJSW between weight-bearing (WB) and NWB images and the association with mediolateral (ML) knee stability.

Methods

23 TKAs from an ongoing RSA study were included. At one-year follow-up, WB and NWB RSA examinations were obtained and the ML stability was evaluated. For each examination the mJSW and femoral-tibial contact locations were measured. A linear regression model was used to analyze the association between the mJSW difference (NWB – WB) with the ML stability and contact locations.

Results

The mean mJSW difference was 0.28 mm medially and 0.20 mm laterally. 4 TKAs had medium (5 - 9 deg) and 19 TKAs had high (< 5 deg) ML stability. A higher mJSW difference was found for TKAs with medium stability (0.36 mm, $p = 0.01$).

Conclusion

In conclusion mJSW measurements in existing (NWB) RSA studies are influenced by knee laxity, but may still provide information on wear progression based on TKA with high ML stability. Because of the difference in contact point locations between WB and NWB positioning and the resulting difference in mJSW, a direct comparison between mJSW from WB and NWB data is not possible.

4-1 Introduction

Polyethylene (PE) wear is a leading cause for failure of total knee arthroplasties (TKAs) [16, 20, 22]. The impact of PE wear is expected to increase further as the incidence of TKAs increases because of our aging and increasingly obese population [5, 11, 54]. In addition, TKAs are applied more often in younger patients that have a more active lifestyle than older patients [6, 11].

Currently, the PE wear of new implant designs or implant materials is evaluated with in vitro knee simulator studies before market introduction [66-68]. These studies do not incorporate the effect of patient specific and surgery specific factors to PE wear [69]. This can lead to unforeseen complications. Alternatively, PE wear can be assessed *in vivo* by measuring the minimum joint space width in radiographs. However, studies using these measurements are uncommon, which may be related to the low precision of conventional *in vivo* wear measurements. Errors up to several millimeters have been reported and obtaining sufficient power is laborious [28]. For example, to distinguish a difference of 0.2 mm in a clinical study approximately 250 patients would be required (2-sided power calculation, SD = 1 mm, alpha = 0.95).

Model-Based Roentgen Stereophotogrammetric Analysis (MBRSA) is an imaging and analysis technique which is known for its high accuracy in measuring migration of implants, which is used as a predictor for survival of knee prostheses [44, 46, 70]. Several studies showed that techniques such as MBRSA can also be used to measure PE wear based on mJSW assessments [43, 71]. We developed and validated such an mJSW measurement for MBRSA in a previous study [72]. Now, this measurement technique can be applied to previous RSA studies on TKA migration, potentially providing information on wear progression.

In most of these RSA studies, however, images were acquired with a supine, non-weight-bearing patient position whereas the joint is loaded in radiographs that are used for conventional wear measurements. Literature states that weight-bearing positions are required for reliable *in vivo* wear assessments in TKA [27, 28, 73, 74]. In supine position, the femoral and tibial components may partially loose contact (due to laxity of the joint), causing the measurement to differ from the actual insert thickness. This requirement has never been fully validated for knee prostheses, while for hip prostheses no difference in wear measurements between weight bearing and non-weight bearing images was found [75]. In case the measurement can detect PE wear progression in supine RSA images, ample data would become available from a multitude of clinical evaluation studies of TKA where successive supine X-rays were made for other purposes.

The primary aim of this study is therefore to determine whether the mJSW measurement differs between weight-bearing and non-weight-bearing positions. A secondary aim is to determine whether this mJSW difference can be related to knee laxity. This is analyzed by comparing TKAs with different mediolateral stability. We hypothesize that a lower mediolateral stability (thus a higher knee laxity) results in larger difference in the mJSW measured in WB and NWB positions.

4-2 Methods

4-2-1 Data

RSA image pairs and knee stability data were analyzed for 23 patients in an ongoing prospective RSA study conducted in 'het Langeland Ziekenhuis' (Zoetermeer, the Netherlands). All patients received a Stryker Triathlon Posterior Stabilized (PS) fixed bearing knee prosthesis. The cohort consisted of 7 males (30%) and 16 females (70%) and aged between 50 and 83 years (median 63 years). All patients gave informed consent to participate in this study.

At the one-year follow-up evaluation, RSA examinations were done in both a standing, weight-bearing (WB) and supine, non-weight-bearing (NWB) patient position. The mediolateral (ML) stability of the TKA was evaluated in degrees. TKAs were classified as having high stability (< 5 deg) or medium stability (5-9 deg). For the RSA examination in supine position a calibration box (Carbon box, RSA Core, dep orthopaedics, LUMC , the Netherlands) was mounted beneath the examination table [32]. For the examination in standing position, this calibration box was positioned vertically. The stereo images were acquired digitally using one mobile X-ray system (Siemens Mobilette, Siemens AG, Munich, Germany) and one in-room X-ray system (didi-series, Philips, Eindhoven, The Netherlands). The images had a pixel spacing of 0.2 by 0.2 mm.

All RSA analyses were carried out at the Leiden University Medical Center (dept Orthopeadics). 2D/3D registration was applied to the stereo images to reconstruct the position and orientation of the femoral and tibial components. This registration was done with model-based RSA software (RSAcore, dep orthopaedics, LUMC , the Netherlands) based on a standardized RSA analysis. This analysis consists of the consecutive steps of image calibration, edge detection and 2D-3D registration based on triangulated surface models using edge matching[17, 76]. For the femoral component computer aided design (CAD) models were used which were obtained from the manufacturer. The tibial models were reverse engineered (RE).

4-2-2 mJSW measurement

For all 46 RSA examinations ((WB + NWB) x 23 TKAs) the mJSW and the contact location of the medial and lateral condyles were measured. The mJSW is defined as the minimum distance between the metal tibial tray and the femoral condyle. The contact location is expressed in tibial tray coordinates x_{AP} and x_{ML} (Figure 4-10).

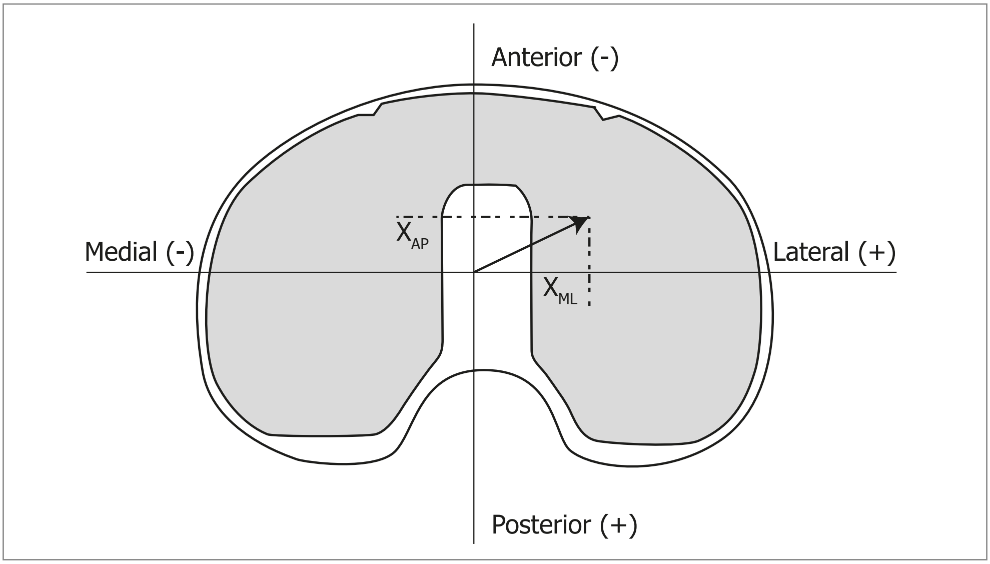


Figure 4-10. Top view of a tibial component with the coordinate system used to specify the contact location with respect to the tibial tray. The origin of the system coincides with the center of the bounding box of the insert. An example vector x is drawn with the AP and ML components indicated.

4-2-3 Analyses

We calculated the difference in mJSW and contact location coordinates (d_{AP} and d_{ML}) between each weight-bearing and non-weight-bearing examination (NWB – WB). For d_{AP} this difference was calculated based on the absolute AP coordinates (|NWB| - |WB|), i.e. the difference in distances from the AP axis.

A linear regression model was used to analyze the mJSW difference and its association with the ML stability of the TKA. In this model the variables d_{AP} and d_{ML} were used as covariates. The rationale for using them is that the insert surface of the Triathlon total knee is not flat and therefore a difference in contact locations is also related to a difference in mJSW measured. This adds variability, which is not related to knee laxity, and therefore adding the variables d_{AP} and d_{ML} gives a better distinction between these effects.

4-3 Results

Table 4-6 shows the descriptive statistics of the WB and NWB mJSW values and their difference. On average, there was a positive difference over the 23 TKAs (0.28 mm medial and 0.20 mm lateral), meaning that a larger mJSW was measured in NWB position. The standard deviation of the mJSW difference was 0.54 mm medially and 0.47 mm laterally. The standard deviations of the WB and NWB mJSW values were larger than that of the mJSW difference, because the former include inter-patient variability of the inserted liner thickness, which can be either 8 mm, 10 mm or 12 mm for the Triathlon total knee.

Table 4-6. The means, standard deviations (SD) and ranges of the WB and NWB mJSW measurements and their difference (diff). All values are expressed in millimeters.

		WB	**NWB**	**diff**
Medial (N = 23)	Mean	8,03	8,31	0,28
	SD	1,77	1,80	0,54
	[min-max]	[5.89 – 13.29]	[5.90 – 12.98]	[-0.49 –1.81]
Lateral (N = 23)	Mean	8,40	8,60	0,20
	SD	1,82	1,87	0,47
	[min-max]	[6.05 – 13.03]	[6.12 – 13.80]	[-0.93 – 0.89]

The distributions of the contact locations for the WB and NWB positions are displayed in Figure 4-11. For the medial condyle, the WB and NWB distributions are very similar, albeit the NWB distribution was 2.5 mm more anterior on average (*t*-test, $p = 0.02$) and had a larger variation in the direction of the ML coordinate.

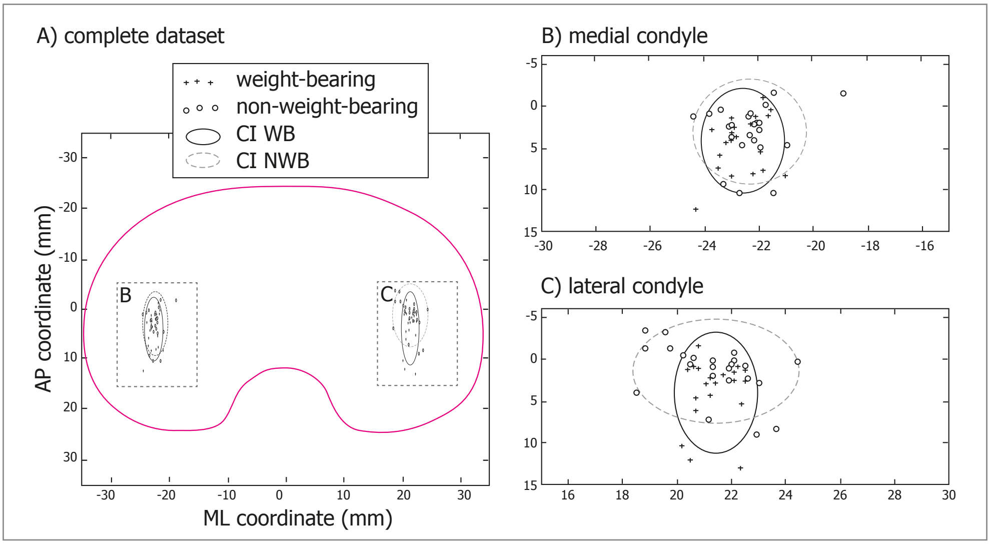


Figure 4-11. Scatterplots of the mJSW coordinates of the examinations for all TKAs and both WB and NWB datasets. (A) overview with the complete dataset and the perimeter of a size 4 triathlon tibia component as reference. On the right, zoomed in views of the medial (B) and lateral (C) datasets are shown. The ellipses indicate the 95% confidence interval of the data (1.96/std dev in the AP and ML directions).

Table 4-7. Summary of the linear regression model. Values and confidence intervals (95% CI) of the coefficients in the linear regression model. Values are expressed in mm mJSW difference per 1 unit change in the coefficient. The dependent variable is the difference in mJSW between NWB and WB positions. Coefficient values printed in bold are statistically significant ($p < 0.05$, Wald Chi Square)

	Medial condyle			**lateral condyle**		
Coefficient	**Value**	**Sig.**	**95% CI**	**Value**	**Sig.**	**95% CI**
Intercept	.069 mm	0.16	[-0.05 – +0.18]	.096	0.10	[-0.02 – +0.22]
d_{ML}	-.368 mm/mm	0.00	[-0.45 – -0.28]	-.265	0.00	[-0.33 – -0.20]
d_{AP}	-.040 mm/mm	0.01	[-0.08 – -0.00]	-.029	0.03	[-0.06 – +0.00]
ML_stab = 5-9deg	.362 mm	0.00	[+0.10 – +0.62]	.071	0.50	[-0.17 – +0.31]

Table 4-7 shows the results of the regression analysis. The analysis showed that ML stability has a significant influence on the mJSW difference for the medial condyle. Four TKAs had a medium stability (5-9 deg) and 19 TKAs had a high ML stability (<5 deg). On average, TKAs with a medium stability had a 0.362 mm larger mJSW difference (p=0.01) compared to TKAs with a high stability. This finding confirms the hypothesis that the difference in mJSW is related to knee laxity. No significant correlation or interactions were found between any of the coefficients d_{ML}, d_{AP} and ML stability (Pearson's correlation and t-tests).

A strong association was found between the mJSW difference and the difference in contact location (d_{ML} and d_{AP}). The coefficient for d_{ML} had the largest magnitude and has the following meaning: a 1 mm shift in contact location further from the center AP axis in the NWB position with respect to the WB position relates to a change in mJSW difference of -0.368 mm laterally and -0.265 mm medially.

To display the effect of ML stability, Figure 4-12 shows a scatterplot of the mJSW difference both for the original data (A) and the residual data (B) after the effects of the covariates d_{ML} and d_{AP} are corrected. As can be seen, in the residual data the variance is much lower and more consistent over the subgroups. In addition, the difference between the stability groups for the medial condyle is distinguishable.

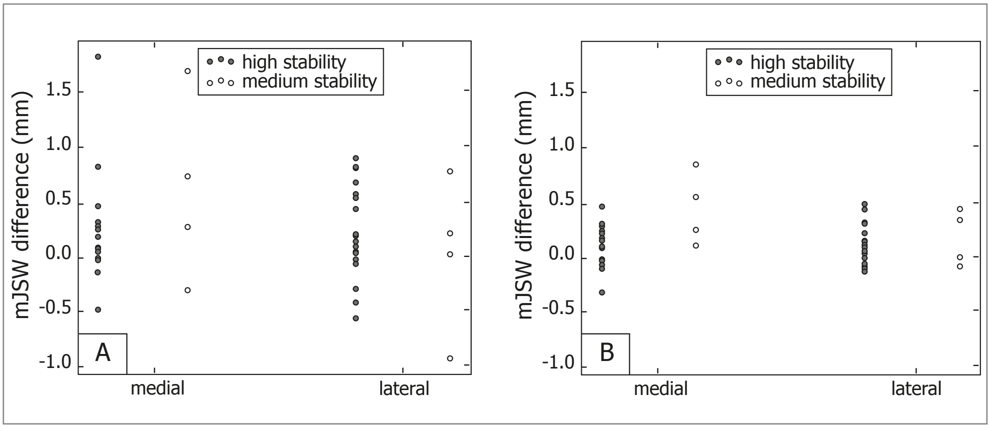


Figure 4-12. Comparison of the mJSW difference between ML stability groups (A, original mJSW differences; B, mJSW differences corrected for d_{AP} and d_{ML}, based on the linear regression model).

4-4 Discussion

We found that the mJSW measured in non-weight-bearing position was larger than the mJSW measured in weight-bearing position. The mean difference for the 23 TKAs was 0.28 mm and 0.20 mm for the medial and lateral condyle respectively. This difference can be caused by both knee laxity, but also by a difference in the measurement location between the WN and NWB positions.

The regression analysis showed that the knee stability was strongly associated with the mJSW difference for the medial condyle. TKAs with a medium stability had a 0.36 mm higher mJSW difference compared to TKAs with a high stability. This effect can be seen in the residual plot of the mJSW differences (Figure 4-11) and was statistically significant in the linear regression model (Wald's Chi Square, $p = 0.001$).

This finding supports our proposition that the mJSW measured in non-weight-bearing patient position is influenced by knee laxity and this limits the applicability of the measurement to assess wear progression in previous longitudinal RSA studies. As PE wear is related to TKA stability, more prostheses may become unstable during follow-up and cannot be measured, leading to an unacceptable selection bias. Nonetheless, during the initial stage of PE wear progression (as long as a TKA remains stable) the mJSW loss could still be detectable and the measurement can be used as an early predictor. Since data on the TKA stability is available in most RSA studies, this is a topic worth further investigation.

We only found a significant effect of ML stability for the medial condyle. This can be explained by the way prostheses are implanted and wear progression occurs. During knee replacement surgery ligament abnormalities are balanced to provide stability to the prosthesis. As wear of total knee prostheses is dominant at the medial condyle, due to the adduction moments at the knee during walking, a relative instability is expected to occur most frequently at this condyle [77].

A limitation of the study is that only four TKAs were included with medium stability, which adds constraints to the conclusions concerning the effect of ML stability. In that regard, it is interesting to notice that the effect sizes were different per condyle

(0.36 mm medially vs. 0.071 mm laterally). This difference can be coincidental, as the confidence intervals of the medial and lateral effect sizes overlap. It is also possible that this difference is related to the kinematics of the knee prosthesis, i.e. knee laxity due to ML stability could differ per condyle.

We assumed that the relation between the mJSW difference and the difference in contact location between the weight-bearing and non-weight-bearing positions was linear, in order to separate the effects of physical difference in insert thickness and joint separation. This assumption seems correct because of the strong associations that were found in the regression model and that no covariation or interactions were found between the variables. Still, if the joint separation size is also related to the contact location, then its effect size can be suppressed by the linear model and an exact model of the insert height profile should be used if a more accurate description is required of this effect.

In conclusion, the insert thickness measurement in non-weight-bearing positions is compromised if the TKA is unstable, and should not be used in those cases. As no significant difference in mJSW was found between WB and NWB positions in TKAs with high stability, the mJSW measurement may still reveal wear trends based on NWB data if cases are carefully selected. The relation between increasing wear, its effect on stability of the TKA and the effect on the accuracy of mJSW measurements should be studied further before they can be used for wear assessment with existing RSA studies.

References

[5] Leskinen, J., et al., *The incidence of knee arthrolasty for primary osteoarthritis grows rapidly among baby-boomers - a population-based study.* Arthritis & Rheumatism, 2012. **64**: p. 423-428.

[6] Kurtz, S.M., et al., *Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030.* Clinical Orthopaedics and Related Research, 2009. **467**: p. 2606-12.

[11] Carr, A.J., et al., *Knee replacement.* The Lancet, 2012. **379**: p. 1331-1340.

[16] Naudie, D.D.R., et al., *Wear and osteolysis around total knee arthroplasty.* The Journal of the American Academy of Orthopaedic Surgeons, 2007. **15**: p. 53-64.

[17] Schmalzried, T.P. and J.J. Callaghan, *Current Concepts Review-Wear in Total Hip and Knee Replacements.* The Journal of Bone & Joint Surgery, 1999. **81**(1): p. 115-136.

[20] Sundfeldt, M., et al., *Aseptic loosening, not only a question of wear: a review of different theories.* Acta Orthopaedica, 2006. **77**: p. 177-97.

[22] Sharkey, P.F., et al., *Why are total knee arthroplasties failing today?* Clinical Orthopaedics and Related Research, 2002. **404**: p. 7-13.

[27] Miller, T.T., *Imaging of knee arthroplasty.* European Journal of Radiology, 2005. **54**: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, *Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses.* The Journal of Arthroplasty, 2003. **18**: p. 860-866.

[32] Kaptein, B.L., et al., *A new model-based RSA method validated using CAD models and models from reversed engineering.* Journal of Biomechanics, 2003. **36**: p. 873-882.

[43] Short, A., et al., *A novel method for in vivo knee prosthesis wear measurement.* Journal of Biomechanics, 2005. **38**: p. 315-322.

[44] Garling, E.H., E.R. Valstar, and R.G.H.H. Nelissen, *Comparison of micromotion in mobile bearing and posterior stabilized total knee prostheses: a randomized RSA study of 40 knees followed for 2 years.* Acta Orthopaedica, 2005. **76**: p. 353-61.

[46] Nilsson, K.G. and J. Kärrholm, *RSA in the assessment of aseptic loosening.* The Journal of Bone & Joint Surgery. British Volume, 1996. **78**: p. 1-3.

[54] Robertsson, O., et al., *Past Incidence and future demands for knee arthroplasty in Sweden.* Acta Orthopaedica, 2000. **71**: p. 376-380.

[66] Chapman-Sheath, P.J., et al., *In vitro assessment of proximal polyethylene contact surface areas and stresses in mobile bearing knees.* Medical Engineering & Physics, 2003. **25**: p. 437-443.

[67] McEwen, H.M.J., et al., *The influence of design, materials and kinematics on the in vitro wear of total knee replacements.* Journal of Biomechanics, 2005. **38**: p. 357-65.

[68] Fregly, B.J., et al., *Computational wear prediction of a total knee replacement from in vivo kinematics.* Journal of Biomechanics, 2005. **38**: p. 305-14.

[69] Collier, M.B., et al., *Factors associated with the loss of thickness of polyethylene tibial bearings after knee arthroplasty.* The Journal of Bone & Joint Surgery. American Volume, 2007. **89**: p. 1306-14.

[70] Nouta, K.A., et al., *All-polyethylene tibial components are equal to metal-backed components: systematic review and meta-regression.* Clinical Orthopaedics and Related Research, 2012. **470**(12): p. 3549-59.

[71] Blagojevic, M., et al., *Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis.* Osteoarthritis and Cartilage, 2010. **18**(1): p. 24-33.

[72] van IJsseldijk, E.A., et al., *The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA.* Journal of Biomechanics, 2011. **44**: p. 2724-7.

[73] Knight, J.L., R.D. Atwater, and J. Guo, *Early failure of the porous coated anatomic cemented unicompartmental knee arthroplasty. Aids to diagnosis and revision.* The Journal of Arthroplasty, 1997. **12**: p. 11-20.

[74] Buckland-Wright, C., *Which radiographic techniques should we use for research and clinical practice?* Best Practice & Research in Clinical Rheumatology, 2006. **20**: p. 39-55.

[75] Bragdon, C.R., et al., *Standing versus supine radiographs in RSA evaluation of femoral head penetration.* Clinical Orthopaedics and Related Research, 2006. **448**: p. 46-51.

[76] van IJsseldijk, E.A., et al., *A model-based approach to measure the minimum joint space width of total knee replacements in standard radiographs.* Journal of Biomechanics, 2012. **45**: p. 2171-5.

[77] Delport, H., et al., *A kinematic comparison of fixed-and mobile-bearing knee replacements.* The Journal of Bone & Joint Surgery. British Volume, 2006. **88**(8): p. 1016-1021.

5

Validation of a model-based measurement of the minimum insert thickness of total knee arthroplasty

Validation of a model-based measurement of the minimum insert thickness of total knee arthroplasty

E.A. van IJsseldijk[a], M.K. Harman[b], J. Luetzner[c], E.R. Valstar[a], B.C. Stoel[d], R.G.H.H. Nelissen[a], B.L. Kaptein[a]

[a] *Biomechanics and Imaging Group, Department of Orthopaedics, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*

[b] *Clemson University, Department of Bioengineering, 301 Rhodes Engineering Research Center, Clemson, 29634, USA.*

[c] *University Hospital Carl Gustav Carus, Department of Orthopaedic and Trauma Surgery, TU Dresden, Fetscherstrasse 74, 1307 Dresden, Germany.*

[d] *Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands*

Published as "Validation of a model-based measurement of the minimum insert thickness of total knee prostheses", Bone and Joint Research, 2014, 3(10), pages 289-296

Abstract

Introduction

Wear of polyethylene inserts plays an important role in failure of total knee replacement and can be monitored in vivo by measuring the minimum joint space width in anteroposterior radiographs. The objective of this retrospective cross-sectional study was to compare the accuracy and precision of a new model-based method with the conventional method by analysing the difference between the minimum joint space width measurements and the actual thickness of retrieved polyethylene tibial inserts.

Method

Before revision, the minimum joint space width values and their locations on the insert were measured in 15 fully weight-bearing radiographs. These measurements were compared with the actual minimum thickness values and locations of the retrieved tibial inserts after revision.

Results

The mean error in the model-based minimum joint space width measurement was significantly smaller than the conventional method for medial condyles (0.50 vs 0.94 mm, $p < 0.01$) and for lateral condyles (0.06 vs 0.34 mm, $p = 0.02$). The precision (standard deviation of the error) of the methods was similar (0.84 vs 0.79 mm medially and both 0.46 mm laterally). The distance between the true minimum joint space width locations and the locations from the model-based measurements was less than 10 mm in the medial direction in 12 cases and less in the lateral direction in 13 cases.

Conclusion

The model-based minimum joint space width measurement method is more accurate than the conventional measurement with the same precision.

5-1 Introduction

Polyethylene is used as bearing material in total knee arthroplasties (TKAs) and its wear plays an important role in TKA failure [22]. Remarkably, standardized (computer assisted) tools for the *in vivo* assessment of polyethylene wear in TKA do not exist. Rather, planar radiographs are the medical standard for routine monitoring of TKA performance and they are used to estimate changes in polyethylene insert thickness during clinical follow-up. This thickness is quantified with the minimum joint space width (mJSW), which is the apparent distance between the metal tibial tray and the femoral condyles in standard frontal plane radiographs [27, 29, 30]. The insert thickness and its change over time can predict TKA failure [23, 24]. However, the conventional mJSW method is applied to image projections, which is subject to parallax errors that occur when the metal tibial baseplate surface is not aligned with the X-ray beam during sequential radiographic assessments. mJSW measurement errors of up to 2 mm are not exceptional and numerous follow-up visits are required to obtain a reliable estimation of the wear rate [28, 29].

In our earlier work a novel, model-based method was presented to measure the mJSW in standard anterioposterior radiographs using highly accurate model-based roentgen stereophotogrammetric analysis (RSA) software [72, 76]. This method has two advantages over the standard mJSW measurements: the effect of parallax errors is reduced by applying a 3-dimensional reconstruction of the prosthesis components using surface models and it gives insight into both the magnitude and location of the mJSW. For a fixed bearing prosthesis, *in vitro* validation showed that the model-based method is superior in accuracy (mean = -0.03 mm vs. 0.20 mm), precision (Standard Deviation = 0.19 mm vs. 0.40 mm) and absolute error (mean = 0.14 mm vs. 0.35 mm) compared with the conventional method [76]. Thus, this method has the potential to improve the accuracy of mJSW measurements, enabling more accurate detection of wear-related complications and improving the power of clinical studies evaluating differences in wear rates between different TKA designs.

In this retrospective cross-sectional study the actual thickness of retrieved polyethylene tibial inserts was compared with the mJSW measurements acquired using the model-based and conventional methods applied to weight-bearing pre-revision radiographs. The primary objective is to compare the accuracy and precision of these mJSW measurement methods using the insert thickness measured from TKA retrievals as a “gold standard”. The secondary objective is to investigate whether the mJSW location determined in the model-based method corresponds to wear locations evident on the explanted polyethylene inserts.

5-2 Method

5-2-1 Data

We searched a database of explanted TKAs catalogued in an Implant Retrieval Program previously established with institutional review board approval (clinical protocol number in Germany EK348112009; retrieval analysis protocol number in USA IBC2011-26) and patient informed consent. Wear scars on polyethylene tibial inserts of 60 fixed-bearing TKAs retrieved from a single clinic (University Hospital Carl Gustav Carus, Dresden) were grossly assessed using optical microscopy to visualize the damage modes and physical touch to detect changes in the articular surface contour. Fifteen posterior cruciate ligament retaining TKAs ultimately were selected to represent a wide range of articular wear scar sizes and shapes, ensuring that the validation study was meaningful for the extensive wear scar variations that can occur in clinical practice [78].

Table 5-8 lists clinical information such as the TKA design, duration of *in vivo* TKA function, the reasons for revision surgery and the grade of the wear scar (mild, moderate or severe). Wear scars were graded as mild if the damage modes visibly disrupted the machine marks on the articular surface without causing a perceptible change in the articular geometry (6 TKAs); moderate if the damage modes visibly disrupted the machine marks on the articular surface and the wear scar was tangible when physically touching the articular surface (5 TKAs); and severe if there was visibly gross material loss (e.g. delamination) and a notable tactile change in the articular geometry due to gross disruption of the bearing surface (4 TKAs).

For each TKA, the most recent anteroposterior planar radiograph was selected from those acquired during routine clinical exams prior to the revision surgery. The radiographs were acquired with a Siemens Aristos FX Axiom imaging device (0.143 mm per pixel). All patients were instructed to remain fully weight-bearing on both limbs. The selected radiographs include unilateral (n=11) and bilateral (n=4) exposures. The radiographs were transmitted in DICOM format following a de-identification process to protect patient privacy in preparation for the radiographic assessments.

Individual 3-D surface models (triangulated meshes) of the explanted components (metal tibial baseplate, polyethylene tibial insert, metal femoral component) were generated using reverse engineering software and a 3-D laser scanner (Next Engine, Santa Monica, CA, USA). These scans had an accuracy of 0.1 mm.

Table 5-8. Description of the fifteen TKAs used in this study.

Case	TKA Design	Wear degree	TKA Lifetime (months)	Lifetime after radiograph[b] (months)	Reason for revision
K2004	TC-Plus	Mild	41	0.7	Infection
K2133	TC-Plus	Mild	17	0.2	Pain
K2145	TC-Plus	Mild	24	3.0	Infection
K2154	Zimmer NK	Mild	50	12.1	Infection
K2171	TC-Plus	Mild	34	0.0	Painful flexion, infection
K2178	TC-Plus	Mild	19	2.5	Infection
K2035[a]	TC-Plus	Moderate	23	0.0	Infection
K2132	TC-Plus	Moderate	86	0.2	Infection
K2137	TC-Plus	Moderate	130	23.6	Suspected osteolysis later diagnosed as metastasis
K2144	TC-Plus	Moderate	132	3.1	Aseptic loosening
K2175	TC-Plus	Moderate	60	0.0	Infection
K2046[a]	Encore Foundation	Severe	144	4.0	Aseptic loosening
K2156[a]	Stryker 7000	Severe	77	0.2	Infection
K2159[a]	Sulzer Protek	Severe	108	1.6	Infection
K2161	TC-Plus	Severe	108	0.2	Infection

[a] For these TKAs, double-leg standing radiographs were used for measuring mJSW, for all other TKA a single-leg standing radiograph was used.

[b] The period between the radiograph acquisition and revision surgery

List of manufacturers: TC-Plus (Smith & Nephew, London, UK); Zimmer NK (Zimmer, Warsaw IN, USA); Encore Foundation (DJO surgical, Vista CA, USA); Stryker 7000 (Stryker, Kalamazoo MI, USA); Sulzer Protek (Protek Medical Product Inc., Coralville, IA, USA).

5-2-2 Assessment methods

The mJSW was measured on the pre-revision radiograph using both the conventional (C) and Model-based (MB) methods; the true insert thickness (d_0) and position (p_0) on the medial and lateral compartments were measured from the scanned models of the polyethylene inserts. The details of these assessments are described below and depicted in Figure 5-13. Last, the articular wear scar on the insert was identified by digitizing the periphery of the worn area.

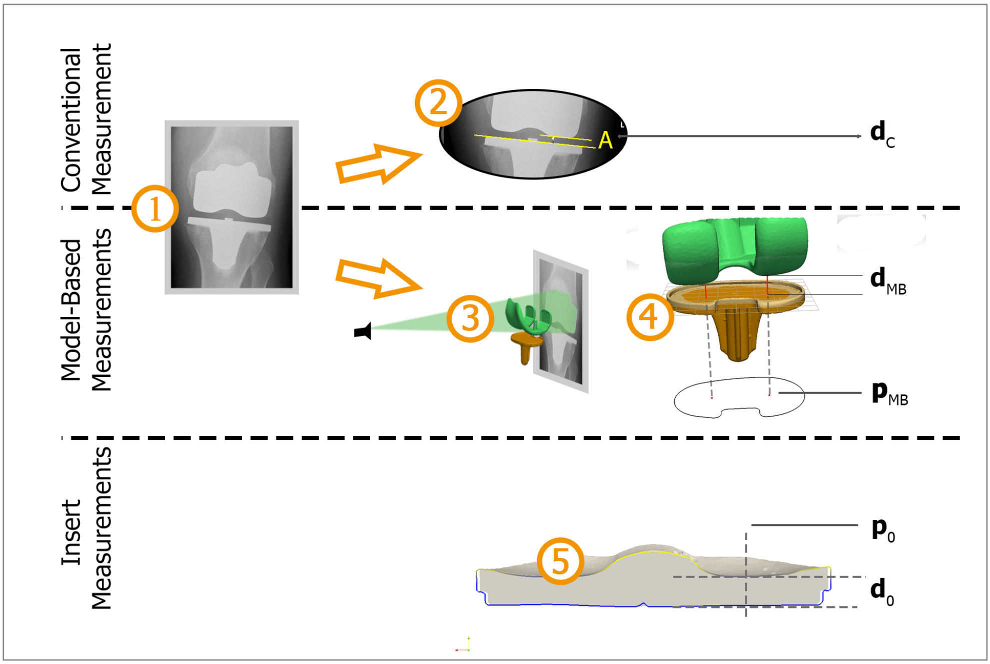


Figure 5-13. Overview of the measurement methods applied for a single total knee replacement (TKR). The rows in the figure represent the measurement methods that were compared: 1) the input radiograph; 2) the conventional insert thickness measurement; 3) 2D/3D matching of the component models; 4) model-based mJSW measurement; 5) the minimum insert thickness and location based on the 3D laser scan of the insert.

Conventional mJSW method

In the conventional mJSW method, the insert thickness (d_C) was assessed directly in the radiographic image, based on the *metal-to-middle* method [28]. This assessment was conducted by an experienced orthopaedic surgeon and an experienced researcher (HvdL & EvIJ) and the average values of the observations were used in the further analysis. Commercially available software was used (Digimizer, MedCalc software, Mariakerke, Belgium) for annotation, image processing and measurement of distances. A reference line was drawn that annotated the superior rim of the metal tibial baseplate at its largest medial-lateral width. The shortest, tibiofemoral distances between this line and the distal femoral condylar edges were measured. The tibial component rim is used for the capture mechanisms securing the polyethylene tibial inserts to the metal baseplate. Therefore, the height of the rim above the tibial baseplate surface was measured by one observer (EvlJ) at three locations using Magics (Materialize, Leuven, Belgium). The mean height was added to the tibiofemoral distances, yielding the final estimate of the insert thickness.

Image magnification was calculated using the ratio between the tibial tray widths in the image silhouette and in the scanned model. This was used to convert all image-based mJSW measurements to real-world dimensions, recorded as the medial d_c and lateral d_c.

Model-based mJSW method

In the model-based method, the mJSW (d_{MB}) was assessed using triangulated surface models of the components (tibia, insert and femur) and using model-based RSA software (Version 3.34, RSAcore, Leiden, The Netherlands)[32]. The tibial model and the insert model were aligned in such a way that the insert's inferior surface and the tibial baseplate's superior surface coincided with the 0xz-plane of the model coordinate systems.

Assessment of the TKAs was initiated with an image-focus calibration step. The pixel size was obtained from the DICOM data and the focus position was set at a 115cm distance to the center of the image, in accordance to the hospital's imaging protocol. Next the tibial and femoral models were matched with the radiographs using 2D-image/3D-model registration.
The mJSW was measured by detecting the femoral condylar model with the shortest distance to the tibial baseplate (d_{MB}). The projection of the points (p_{MB}) was stored and expressed in anterio-posterior (AP) and medio-lateral (ML) coordinates with respect to the center of the tibial baseplate. The measurement was repeated by two researchers (EvIJ and BLK), who independently conducted the registration

and measurement processes. The average values of the observations were used in further analysis.

Insert measurements

Using the 3D laser scan of the explanted polyethylene inserts, the minimum insert thickness in millimeters (d_0) was measured as the minimum perpendicular distance between the inferior backside surface and the articular surface of the insert. The scans were aligned with the tibia models and the locations of the minimum insert thickness (p_0) were expressed in the same coordinate system as in the model-based mJSW method.

One experienced observer (MKH) analyzed the wear scar area of the inserts using the following approach: The wear scar areas were visually identified using an optical stereomicroscope (model Z30L, Cambridge Instruments, Cambridge, MA, USA). Subsequently, the circumference of the insert periphery and the circumference of the wear scars were digitized on calibrated digital images of the articular surface using published photogrammetry methods[79, 80]. The insert circumference was used to map these data to the tibia model coordinate system.

5-2-3 Statistical Analysis

The values Δ_C and Δ_{MB} were calculated as the difference between the respective mJSW assessment d_C and d_{MB} and the reference insert thickness d_0 ($\Delta_C = d_C - d_0$, $\Delta_{MB} = d_{MB} - d_0$). The mean and standard deviation of these differences over the 15 cases were calculated and compared (paired *t*-test). In addition, the mean measurement errors were calculated as the mean of the absolute difference $|_C|$and $|_{MB}|$ and the number of cases having an absolute difference smaller than 1 mm was counted, similar to the analysis by Collier et al[28]. Inter-observer agreement was analyzed with the limits of agreement and Bland-Altman plots per condyle and mJSW measurement method[81].

To investigate whether the model-based mJSW measurement can accurately determine the location of the minimum insert thickness, the locations of the model-based mJSW assessment (p_{MB}) and minimum insert thickness (p_0) were compared. The mJSW accuracy could be associated with the difference these locations and this was tested by computing the correlation between these outcomes. The number of TKAs was counted for which the model-based measurement points (p_{MB}) were within the wear scar periphery.

5-3 Results

5-3-1 mJSW measurements

After enduring functional lifetimes of approximately 1.5 to 12 years, the actual minimum insert thickness measured on these explanted polyethylene bearings ranged from d_0 = 1.99 mm to 7.86 mm medially and 4.97 mm to 7.92 mm laterally (Figure 5-14). The mean difference between the mJSW (d_{MB} or d_C) and insert thickness (d_0) was positive for both methods (Table 5-9), meaning that both methods tended to overestimate the actual minimum insert thickness that was measured from the explanted tibial inserts. The standard deviations of the mJSW measurement methods were similar. The mean measurement error was significantly smaller for the model-based measurement than for the conventional measurement for both the medial condyle (0.50 mm versus 0.94 mm, $p < 0.01$) and lateral condyle (0.06 mm versus 0.34 mm, $p = 0.02$); (paired *t*-tests).

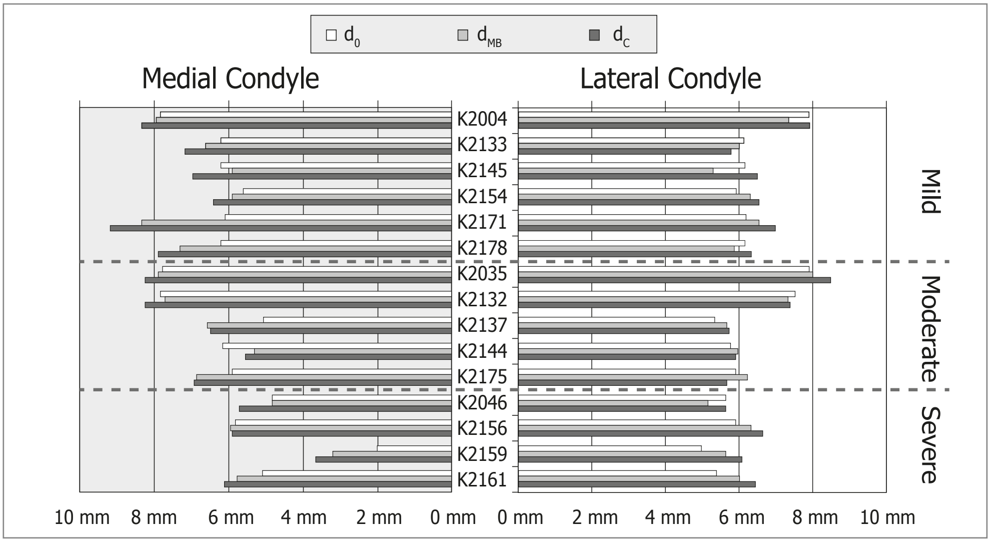


Figure 5-14. Barplots of the estimated insert thicknesses d_C from the conventional mJSW method, d_{MB} from the model-based mJSW method, and actual minimum insert thickness d_0 for each case. The cases are ordered as in Table I and grouped by wear grade.

Table 5-9. Statistics of the differences between the mJSW measurements (conventional D_C and model-based D_{MB}) with respect to the true minimum insert thickness d_0.

	Medial condyle (N=15)			**Lateral condyle (N=15)**		
	Δ_C	Δ_{MB}		Δ_C	Δ_{MB}	
Mean (mm)	0.94	0.50	p = 0.001	0.34	0.06	p = 0.021
Standard deviation (mm)	0.84	0.79	p = 0.772	0.46	0.46	p = 0.982
Mean measurement error (mm)	1.02	0.66	p = 0.001	0.44	0.40	p = 0.311
N(err < 1 mm)	9 (60%)	11 (73%)		13 (87%)	1 5 (100%)	

[1] Paired *t*-test for equal means

[2] Levene's test for homogeneity of variance

The limits of agreement between the observers over the 15 cases were calculated for both mJSW measurement methods. For the model-based mJSW the values were 0.00 ± 0.45 and 0.00 ± 0.54 (mean ± 1.96 * standard deviation) for the medial and lateral condyles respective. For the conventional mJSW these values were -0.22 ± 0.48 and -0.21 ± 0.45. For both condyles a systematic difference was found between the observers for the conventional method (Student *t*-test, $p < 0.01$). The Bland-Altman plots of the outcomes (Figure 5-15) showed no other trends for either mJSW measurement method. Two outliers (K2154 and K2156, both condyles) were found in the distribution of the observer difference for the model-based measurement. For the conventional measurement case a single outlier was found (K2154).

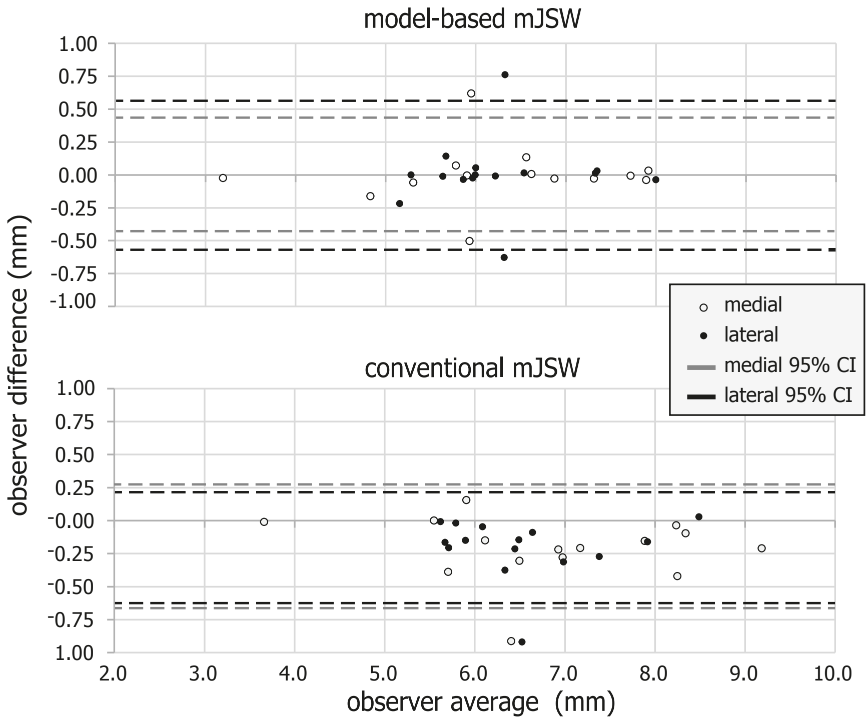


Figure 5-15. **Bland–Altman plots A) of the model-based mJSW and B) of the conventional mJSW method**

5-3-2 Evaluation of the measurement points

The locations of the measurement point (pMB) was compared with the minimum insert thickness location (p0) and the wear scar area (Figure 5-16) and the difference between the points in terms AP and ML distance was computed (Table 5-10). The largest distances were found in the AP direction, where the differences ranged between -18 mm (anterior) and +6 mm (posterior). The Euclidean distance was smaller than 10 mm for 12 out of 15 cases medially and 13 out of 15 cases laterally. The median Euclidian distance was 6 mm and the largest Euclidean distance was 18 mm. For all cases the locations were inside the wear scar area or at the edge of the wear scar area. No significant correlation was found between the Euclidian distance and the measurement error of the model-based mJSW measurement (Spearman's rho= 0.07, P =0.70).

Table 5-10. The differences in position between the femoral contact locations p_{MB} and the minimum insert thickness locations p_0 (as seen in Figure 5-16). Values are expressed in millimeters.

Case	**condition**	**Medial Compartment**		**Lateral Compartment**	
		AP	**ML**	**AP**	**ML**
K2004	Mild	5.98	-1.84	0.63	1.06
K2133	Mild	5.98	-0.22	-2.06	0.81
K2145	Mild	-7.47	3.88	5.05	-0.68
K2154	Mild	0.52	0.76	-7.51	-0.96
K2171	Mild	2.02	-6.38	-1.29	2.26
K2178	Mild	1.67	-1.06	-0.46	1.20
K2035	Moderate	-1.95	-7.26	-0.14	-3.37
K2132	Moderate	-8.97	-8.80	-16.40	-2.59
K2137	Moderate	2.32	-6.57	-8.80	-3.39
K2144	Moderate	-9.04	-2.06	-9.10	-0.17
K2175	Moderate	-12.88	-1.16	-17.76	-2.18
K2046	Severe	-11.02	-3.69	-6.06	3.60
K2156	Severe	0.14	-1.07	-2.51	2.42
K2159	Severe	0.03	-0.53	1.27	1.99
K2161	Severe	4.89	-4.30	1.43	1.78

5-4 Discussion

The primary objective was to compare the accuracy and precision of the model-based mJSW measurement and the conventional mJSW measurement using minimum insert thickness measured from TKA retrievals as a “gold standard”. The accuracy (proximity to the truth) and precision (measurement reproducibility) of both methods was determined by applying the methods to pre-operative radiographs and comparing the outcomes with the minimum thickness of the retrieved inserts. The results showed that the model-based measurement method was more accurate than the conventional method for both condyles (0.50 vs 0.94 mm medially and 0.06 vs 0.34 mm laterally). The precision of the methods was similar (0.84 vs 0.79 mm medially and both 0.46 mm laterally). Both mJSW measurements were more accurate

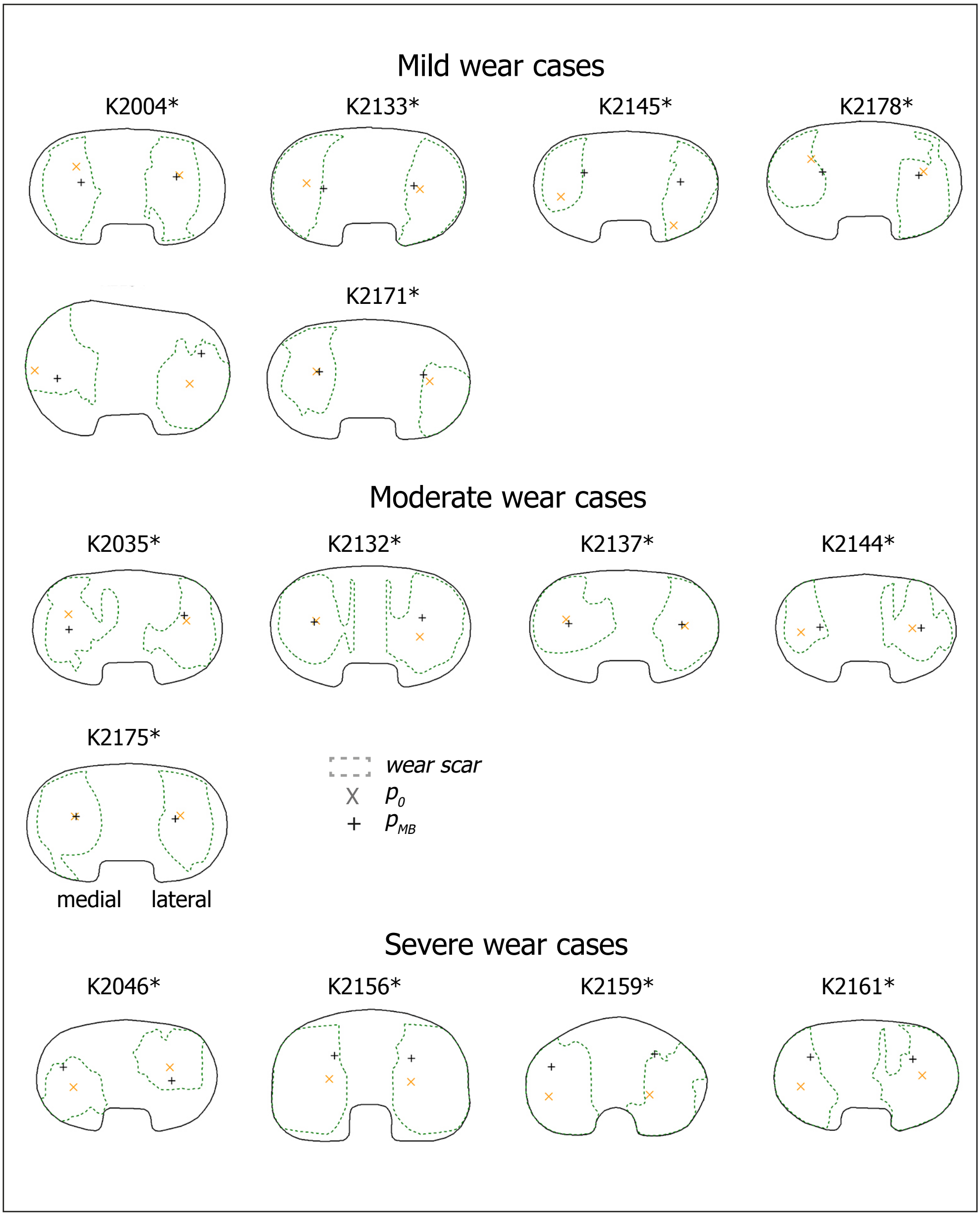


Figure 5-16. Illustrations of the articular surfaces of each explanted insert, showing the wear scar peripheries and locations of the minimum insert thickness (p_0) and the femoral contact (p_{MB}). These illustrations are plotted as looking down on the superior surface of a right knee, with the medial condyle always at the left side of the image. Inserts originating from left TKR are mirrored in this illustration to fit this convention; (indicated with an asterisk (*)).

and precise for the lateral condyle than for the medial condyle. Since this occurred for both methods, this is not a measurement error. Apparently a physical difference existed between the femorotibial distance and the insert thickness, which can be related to various clinical conditions such as varus malalignment.

Concerning the observer reproducibility, the model-based method the mean difference was 0.0 mm and for the conventional method the mean difference between the observers was 0.2 mm. The limits of agreements of the mJSW measurement methods were similar. For the cases K2154 and K2156 a large difference (> 0.5 mm) was found between the model-based observers mJSW measurements. For K2154 some bone cement was still attached to the backside of the tibia baseplate when it was scanned. This introduces a model inaccuracy and complicates the matching procedure, as the respective contours of the tibial metal baseplate should then not be used in the 2D/3D matching. One observer deselected these particular contours, whereas the other observer included this contour part, which may explain the measurement difference. For K2156 one observer did not apply the 2D/3D matching process for the tibia component correctly. This resulted in an out-of-plane positioning error that affected the measurement outcome. Still, the average measurements for these cases were not remarkably far from the actual minimum insert thickness. The outlier for the conventional mJSW measurement (K2154) was related to a difference in setting the height of the reference line at the tibial baseplate.
Four cases stand out (K2137, K2159, K2171 and K2178) as relatively large overestimations (more than 1 mm) of the medial insert thickness for both methods. For K2137 this seems to be related to the image calibration: in the model-based optimization the posterior edge of the femoral component models is approximately 3 cm away from the X-ray detector plate, which is physically unlikely. For K2159 there is a large difference between the measurement location p_{MB} and the actual minimum insert thickness location p_0. For the other cases no obvious explanation could be found, and it may be possible that for these patients there was no actual contact at the mJSW position p_{MB} at the medial side.

Our secondary objective was to investigate whether the explanted inserts truly show wear scars at the points measured by the model-based mJSW technique. The analysis showed that this was true for all inserts. It should be noted that for some cases, such as K2156 and K2132, the wear scar covers the majority of the inserts' articular surface area, which dilutes the information of this observation as any measurements is bound to reside in the wear scar area. Still, this finding supports the proposition that the mJSW measurement is suitable to detect wear.

Concerning the difference between the minimum insert thickness location (p_0) and the femoral contact location (p_{MB}), the findings were volatile. The findings were similar for the medial and lateral condyles: the Euclidean difference was smaller than 10 mm for twelve cases medially and thirteen cases laterally. When this difference was larger than 10 mm, the measurement point always was more anterior than p_0. This could be related to the patient positioning: patients are standing with extended knees during the image acquisition whereas the femoral condyles reposition during dynamic activities [77]. In posterior cruciate ligament retaining TKA, knee flexion during activity can contribute to posterior contact of the femoral condyles and posterior wear scars [82]. This is supported by the observation that three out of four cases with severe wear had a relatively posterior location for p_0. The anteroposterior direction also corresponds to the film-focus direction for a frontal plane radiograph, for which the 2D-3D model matching algorithm is the least accurate. Therefore, the difference in location can also be related to measurement error.

Collier et al. found that conventional mJSW measurements had an accuracy within 1 mm for 82% medially and 58% laterally [28]. This is comparable to the findings with the conventional method in the current study (60% medially and 87% laterally within 1 mm), although the accuracy numbers for the condyles are interchanged. Differences between these results could be caused by the type of prosthesis that was evaluated. Whereas Collier et al. used a single, flat-surfaced Anatomic Modular Knee (Depuy, Warsaw, IN, USA), the measurements in the current study were applied to five different implant designs as to validate our measurement technique as a more generic application to different implant models. This also included designs having a metal rim capture mechanism on the tibial baseplate, which can distort the projection image and for which an alternative approach of the conventional mJSW method had to be used. Moreover, Collier, et al. achieved good measurement accuracy only when TKR were well-aligned relative to the projection plane, necessitating that 28%-39% of their radiographs be discarded from the measurement analysis due to excessive anteroposterior tilt of the tibial baseplate [28, 29]. For the current study all radiographs were utilized regardless of baseplate tilt.

In our prior validation study the model-based mJSW measurement showed a standard deviation of 0.2 mm in case of fixed-bearing TKAs, against 0.79 mm medially and 0.46 laterally in the current study [76]. An explanation for this difference is that repeated measurements for a single TKA were used in the validation study, whereas fifteen different TKAs were measured in our current study. Moreover, in the validation study the inserts were replaced with a flat acrylic block [76]. This approach removed the possibility that sagittal plane curvature of the articular surface could lead to large variations in thickness with only slight deviations in the anteroposterior position of the femoral condyle.

This study was set up in an attempt to capture a representative range of wear severity in a limited number of implant designs and to obtain a first impression of the accuracy that can be obtained with the model-based mJSW method *in vivo*. In future work the data need to be augmented to include a wider range of prosthesis designs with varied insert curvature and to determine the precision of the method when longitudinal data are analyzed.

The model-based mJSW measurement requires accurate tibial and femoral models. In this study, models were generated by reverse engineering prosthesis components that were retrieved from the cohort of included patients. This resulted in the best possible model accuracy for the model-based method [32]. In practice, it will not be possible to use such patient-specific models, as longitudinal assessments of polyethylene wear are conducted without availability of retrieved components. In that case scanned models (reverse engineered models) are recommended, that can be produced based on matching components (i.e. of the same type and size) and the costs of production are relatively low.

Contour detection and optimization can be time-consuming tasks of the model-based mJSW measurement, which might limit the use in clinical evaluation studies. A topic of further research is to reduce the measurement time using further automation of the measurement procedures. Furthermore the measurement could also be improved by reducing the out-of-plane error of the optimization. For example, this could be realized by restricting the freedom of the model pose using prior knowledge on the allowed range of motion of the TKA [83].

In conclusion, the model-based mJSW measurement method delivers a more accurate estimation of the *in vivo* insert thickness from planar radiographs compared with the conventional measurement. In addition, it provides information on the mJSW location, which is indicative for the site of the wear. Further research is required to come to a standardized measurement protocol and to investigate whether the model-based mJSW can hold its accuracy gain in longitudinal data and for a broader range of prosthesis designs.

References

[22] Sharkey, P.F., et al., *Why are total knee arthroplasties failing today?* Clinical Orthopaedics and Related Research, 2002. **404**: p. 7-13.

[23] Engh, C.A., et al., *Radiographically measured total knee wear is constant and predicts failure.* The Journal of Arthroplasty, 2013.

[24] Pijls, B.G., H.M. Van der Linden-Van der Zwaag, and R.G. Nelissen, *Polyethylene thickness is a risk factor for wear necessitating insert exchange.* International Orthopaedics, 2012. **36**(6): p. 1175-80.

[27] Miller, T.T., *Imaging of knee arthroplasty.* European Journal of Radiology, 2005. **54**: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, *Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses.* The Journal of Arthroplasty, 2003. **18**: p. 860-866.

[29] Collier, M.B., et al., *Radiographic assessment of the thickness lost from polyethylene tibial inserts that had been sterilized differently.* The Journal of Bone & Joint Surgery. American Volume, 2008. **90**: p. 1543-52.

[30] Sanzén, L., et al., *Radiographic wear assessment in a total knee prosthesis.* The Journal of Arthroplasty, 1996. **11**: p. 738-42.

[32] Kaptein, B.L., et al., *A new model-based RSA method validated using CAD models and models from reversed engineering.* Journal of Biomechanics, 2003. **36**: p. 873-882.

[72] van IJsseldijk, E.A., et al., *The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA.* Journal of Biomechanics, 2011. **44**: p. 2724-7.

[76] van IJsseldijk, E.A., et al., *A model-based approach to measure the minimum joint space width of total knee replacements in standard radiographs.* Journal of Biomechanics, 2012. **45**: p. 2171-5.

[77] Delport, H., et al., *A kinematic comparison of fixed-and mobile-bearing knee replacements.* The Journal of Bone & Joint Surgery. British Volume, 2006. **88**(8): p. 1016-1021.

[78] Harman, M., et al., *A pictographic atlas for classifying damage modes on polyethylene bearings.* Journal of Materials Science: Materials in Medicine, 2011. **22**(5): p. 1137-1146.

[79] Harman, M.K., et al., *Comparison of polyethylene tibial insert damage from in vivo function and in vitro wear simulation.* Journal of Orthopaedic Research, 2009. **27**(4): p. 540-548.

[80] Harman, M.K., S.A. Banks, and W.A. Hodge, *Backside damage corresponding to articular damage in retrieved tibial polyethylene inserts.* Clinical Orthopaedics and Related Research, 2007. **458**: p. 137-144.

[81] Martin Bland, J. and D. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* The Lancet, 1986. **327**: p. 307-310.

[82] Harman, M.K., S.A. Banks, and W.A. Hodge, *Polyethylene damage and knee kinematics after total knee arthroplasty.* Clinical Orthopaedics and Related Research, 2001. **392**: p. 383-93.

[83] Prins, A., et al., *Detecting femur–insert collisions to improve precision of fluoroscopic knee arthroplasty analysis.* Journal of Biomechanics, 2010. **43**(4): p. 694-700.

6

Validation of the in vivo volumetric wear measurement for total knee prostheses in model-based RSA

Validation of the in vivo volumetric wear measurement for total knee prostheses in model-based RSA

E.A. van IJsseldijk[a], B. Lebel[b], B.C. Stoel[c], E.R. Valstar[a,d], S. Gouzy[b], C. Vielpeau[b], B.L. Kaptein[a]

[a] *Biomechanics and Imaging Group, Department of Orthopaedics, Leiden University Medical Center, The Netherlands*

[b] *Department of Trauma and Orthopaedic Surgery, Centrale Hospital Universitaire Caen, France*

[c] *Department of Radiology, Leiden University Medical Center, The Netherlands*

[d] *Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, The Netherlands*

As published in The Journal of Biomechanics, 2013, 46(7), pages 1387-1391

Abstract

Introduction

Implant failure related to polyethylene wear remains an important issue in total knee arthroplasty. Polyethylene wear is usually assessed in vivo by measuring the remaining insert thickness on X-ray images of the knee. To reflect the amount of wear debris more accurately, a 3-dimensional overlap measurement has been suggested, which is based on implant component models which are matched on calibrated stereo X-ray images using model-based roentgen stereophotogrammetric analysis. The goal of this study was to determine the influence of pose estimation, insert thickness deviation and variation in the tibiofemoral contact location on the accuracy and precision of the measurement using simulations and a phantom experiment.

Results

We found that the pose estimation was the largest source of variation. The 95% prediction interval varied between 111 and 283 mm^3, which is approximately 100–200% of the detected volumetric wear. Insert thickness variation resulted in prediction intervals of 74–174 mm^3. Variation of the tibiofemoral contact location in the phantom experiment gave a prediction interval of 40 mm_3. Large differences in the detected wear volume were found for different flexion angles. At most 56% of the true wear volume was detected (129 of 230 mm^3, 301 of flexion).

Conclusion

In summary, both the accuracy and precision of the volumetric wear measurement were low. The prediction interval of the volumetric wear measurement is at least as large as the measurement outcome itself. This is an important limitation to the applicability of the volumetric wear measurement in clinical practice and further clinical validation is required.

6-1 Introduction

Polyethylene (PE) wear is an important cause of implant failure of total knee arthroplasty (TKA), as it can lead to instability and aseptic loosening [16, 20, 22]. Therefore, an accurate and precise method is required to assess the in vivo progression of PE wear in vivo, which can be used to predict instability and loosening so as to initiate a timely intervention.

The current method to assess the progression of PE wear in vivo is measuring the minimum distance between the femoral condyles and the tibial plateau using radiographic and fluoroscopic imaging [27, 28, 30, 40, 76]. However, this 2-dimensional measurement does not reflect the total volume of wear debris that has been released. Therefore, Gill and coworkers presented a method to measure the in vivo wear volume using 3-dimensional (3-D) geometric models of the implant components, by estimating their 3-D poses (positions and orientations) from stereo X-ray images and calculating the overlap volume with the insert [48].

For the most part the accuracy and precision of this measurement method have not been validated. The goal of this study was to determine the influence of important sources of variation on the accuracy and precision of the volumetric wear measurement. Amongst others, these depend on the 3-D pose estimation and deviations in the original insert thickness as a result of the manufacturing process. Simulation studies were conducted in which the isolated influences of these sources on the measurement were determined.

In practice, wear is often caused by the sliding motion of the femoral component relative to the insert. Therefore, the accuracy and precision of the measurement will also relate to the flexion angle at which the measurement is conducted and the variation in the femoral contact location on the insert. A phantom experiment was done to determine the influence of these sources, using inserts with abrasive wear.

6-2 Materials and Methods

The volumetric wear measurement was conducted based on image pairs that were acquired using a röntgen stereophotogrammetric analysis (RSA) setup with the calibration box in vertical orientation [32]. The image pairs were analyzed with Model-based RSA software (v3.32, Medis Specials, Leiden, The Netherlands) to estimate the poses of the prosthesis components, which are described with triangulated surface models [32]. Since the insert component does not produce clear image contours, its pose was derived from the pose of the tibia model, as they have a fixed spatial relationship.

Volumetric wear was detected by calculating the 3-D overlap region between the femoral and insert component models. A regular 2-D grid was defined (0.8 x 0.8 mm cell size) that coincided with the tibial plateau. For each grid point the overlap distance between the femoral component's surface and the insert surface was calculated. The wear volume was computed using a numerical integration of these distance values based on Simpson's rule.

6-2-1 Simulation Experiments

The influences of pose estimation and insert thickness deviations were determined in simulation experiments. We calculated the difference in the detected volumetric wear as a function of the relative pose of the femoral component with respect to the tibial component. This pose is expressed as $p = (x,y,z,\alpha,\beta,\gamma)^T$, where x, y, and z are the medial, caudal and anterior position parameters and α, β and γ are the corresponding Euler angles (Figure 6-17).

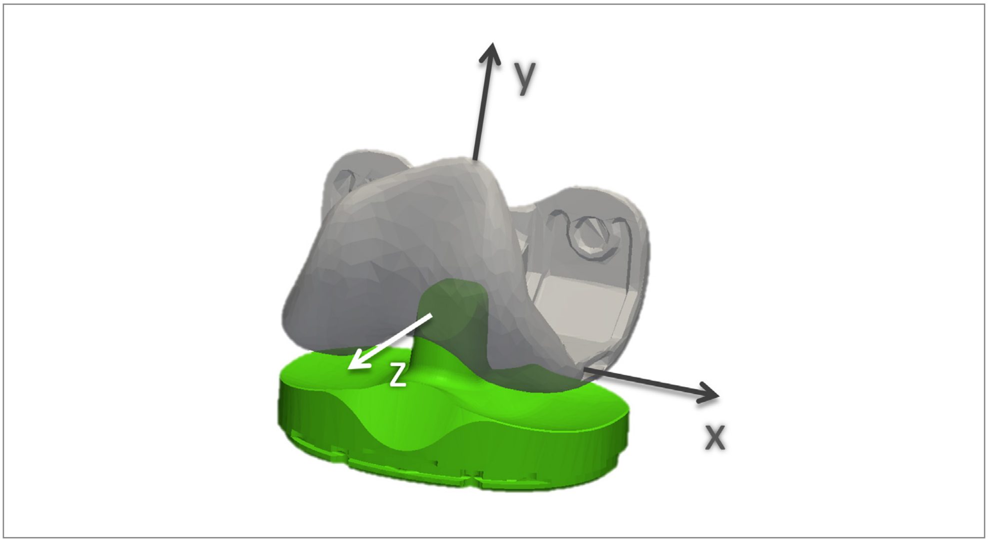


Figure 6-17. The coordinate system that was used in the simulation study.

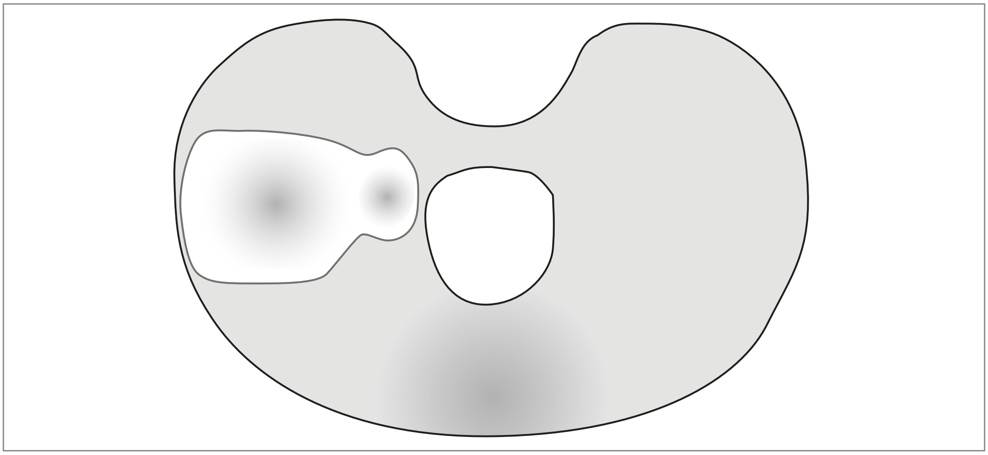


Figure 6-18. Illustration of the predefined wear pool (size = 230 mm³). The shading intensity of the area corresponds to the depth of the wear pool with respect to the insert surface.

The experiments were repeated with eight initial poses $p_{0,j}$ (j=1..8), which were obtained from eight RSA data of patients with size 4 Triathlon PS total knee prostheses (Stryker Europe, Raheen, Ireland).

The effect of pose estimation error was computed in a Monte Carlo Simulation. For each initial pose the detected volumetric wear $w_{0,j}$ was calculated and 500 new poses were generated as $p_{i,j} = p_{0,j} + d_{i,j}$. The pose errors $d_{i,j} = (d_{xi,j}, d_{yi,j}, d_{zi,j}, d_{\alpha i,j}, d_{\beta i,j}, d_{\gamma i,j})^T$ were drawn from a normal distribution with zero mean and a standard deviation (SD) of (0.085 mm, 0.085 mm, 0.22 mm, 0.343°, 0.414°, 0.23°)^T. These SDs were derived from a clinical validation study [33]. For each pose the detected wear volume $w_{i,j}$ and measurement error $e_{i,j} = w_{i,j} - w_{0,j}$ were calculated.

The variation in insert thickness was simulated by varying the caudal position parameter of the relative pose with *D*d, resulting in $p_j = p_{0,j} + (0, \Delta d, 0, 0, 0, 0)^T$. The parameter was varied between +0.12 mm and -0.12 mm, which is the range of the 95% prediction interval assuming that the thickness among insert components of the same type and size vary with an SD of 0.06 mm [28, 84, 85].

Phantom Experiment

The phantom experiment was conducted to assess the influence of variation in the tibiofemoral contact location and the knee angle to the volumetric wear measurement. We used a knee prosthesis (size 4 Triathlon PS) with inserts containing a predefined wear pool and determined how accurately these wear pools could be reconstructed by the volumetric wear measurement.

The wear in the inserts was designed in SolidWorks CAD software (Dassault Systemes, Paris, France). A femoral component model (size 5 Triathlon PS) was placed in bearing contact with the insert model and subsequently moved downward (into the insert). This produced a 3-D overlap volume between the models, which was removed from the insert model. Different sizes and shapes of the wear pool were created (N=6) by varying the flexion angle of the femoral component and the distance over which it was moved into the insert. We used a larger size femur component to simulate wear caused by the sliding motion of the femoral component. The physical insert was manufactured by a computer controlled milling device (Stryker Europe, Raheen, Ireland).

We selected an insert for which the tibiofemoral contact location was consistently found inside the wear pool in the volumetric wear measurement (see Figure 6-18). The data of all other inserts is presented in Appendix 1.

A total knee prosthesis was assembled with the selected insert placed in the tibial component. For analysis and pose estimation 3-D scans of the insert, femoral and tibial components were generated by means of reversed engineering (Introtech, Nuenen, the Netherlands). Based on the insert scan, the shape and volume of the true (predefined) wear pool were determined.

This especially prepared prosthesis was fixed into sawbones. The tibia sawbone was placed in a vertical position on a tripod. The femur sawbone could be positioned on top of the tibia in any flexion angle, as a 7kg balancing weight was used to stabilize the set-up (see Figure 6-19).

The sawbones were placed in a horizontally-oriented RSA imaging setup. Five consecutive RSA image pairs were obtained for three flexion angles (0°, 30° and 60°) resulting into 15 image pairs totally. Before obtaining each of these image pairs, the femoral component was remounted in such a position so that the predefined flexion angle was set (verified by a goniometer) and so that its contact location resided inside the wear pool. By this operation, variation in the tibiofemoral contact location was introduced.

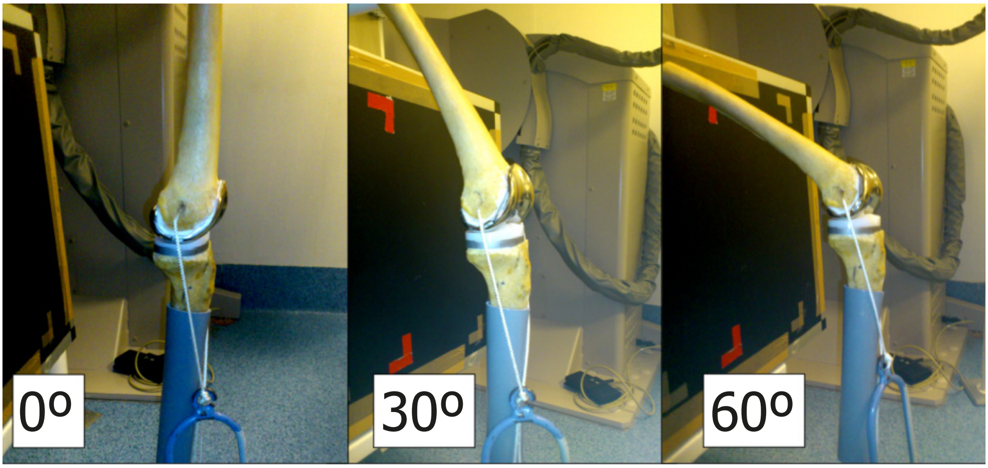


Figure 6-19. The set-up of the phantom experiment during the image acquisition. For each angle of knee flexion an image is shown.

For each RSA image pair, the volumetric wear was assessed and the detected wear pool was compared to the true wear pool, defining both the correctly and falsely detected wear (Figure 6-20). The part of the true wear pool that was not detected was defined as missed wear. The volumes of these quantities were calculated and the means and SDs over the flexion angles were compared.

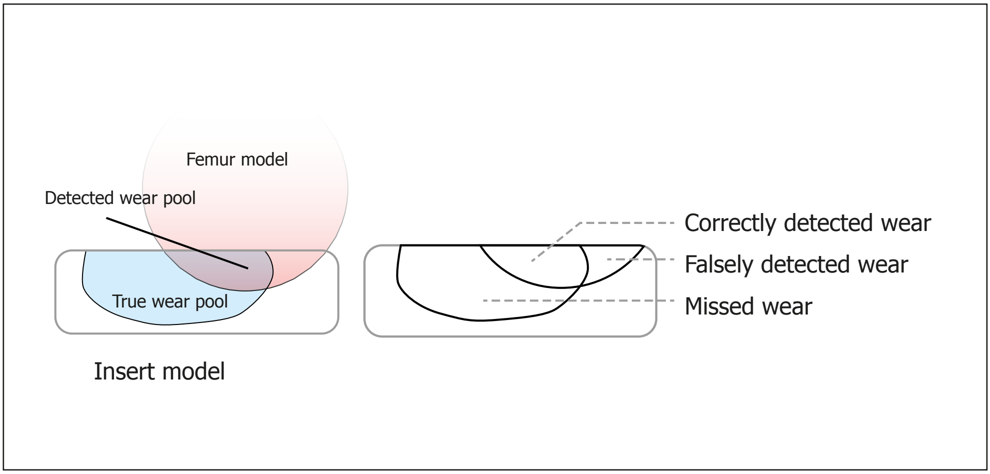


Figure 6-20. Schematic cross-section of an insert with wear and with the femoral-insert overlap measurement. This figure shows how correctly detected wear, falsely detected wear and missed wear are defined.

6-3 Results

6-3-1 Simulation Experiments

Table 6-11 shows the results related to the pose estimation error and the insert thickness variation. For the pose estimation error, the mean wear was slightly larger than w_0 (8 mm^3, $p = 0.001$, paired samples *t*-test). This difference is caused by the non-linear relation between the wear volume detected and the y position. The sizes of the prediction intervals (PI) ranged between 111 and 283 mm^3 and were positively and significantly correlated with w_0 (Pearson's $\rho = 0.96$).

The effect of varying the thickness of the insert (*Dd*) on the detected wear volume can be seen in Figure 6-21. Their relation is not entirely linear, as the size of the slope (measurement error as a function of Dd) declined for increasing *Dd*. The 95% PIs ranged between 74 and 174 mm (Table 6-11).

Table 6-11. Results of the simulations. w_0 is the wear volume corresponding to the initial pose p_0. For the pose estimation error the mean and 95% prediction intervals of the 500 detected wear volumes $w_{i,j}$ are presented. For the insert thickness variation, the 95% prediction intervals of the wear volume are shown, which are defined as the wear volume measured after adding ±0.12 mm to the y-position of the relative pose. The size and relative size of the PIs with respect to the original wear are presented.

	Pose estimation Error				**Insert Thickness Variation**		
original wear (w_0)	mean wear	95% PI	PI size	size / w_0	95% PI	PI size	size / w_0
mm^3	mm^3	mm^3	mm^3	-	mm^3	mm^3	-
51	60	[11 - 122]	111	(2.18)	[21 - 95]	74	(1.45)
67	78	[21 - 154]	133	(1.99)	[29 - 121]	92	(1.37)
84	93	[30 - 178]	148	(1.76)	[42 - 138]	96	(1.14)
125	139	[60 - 235]	175	(1.40)	[74 - 189]	15	(0.92)
157	166	[64 - 282]	218	(1.39)	[95 - 234]	139	(0.89)
163	171	[87 - 273]	186	(1.14)	[114 - 225]	111	(0.68)
172	181	[93 - 290]	197	(1.15)	[110 - 246]	136	(0.79)
304	310	[185 - 468]	283	(0.93)	[222 - 396]	174	(0.57)

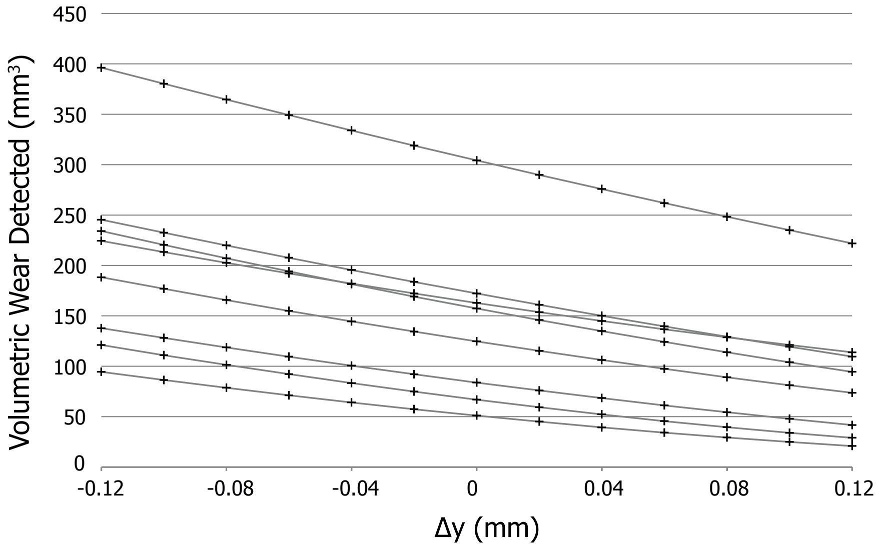


Figure 6-21. The volumetric wear detected as a function of changing the insert thickness (Δd). The eight inputs are presented with separate lines. The limits Δd= ±0.12 mm equal the 95% PI interval of the manufacturing process.

6-3-2 Phantom experiment

The bar plot in Figure 6-22 presents the correctly and falsely detected wear volumes for the 15 image pairs. Below the figure, typical examples of these wear pools are shown per flexion angle. As a reference, the leftmost bar shows the volume of the true wear pool.

A comparison of the results per flexion angle is shown in Table 6-12. The mean of both the correctly and falsely detected wear volumes showed a significant difference between the flexion angles ($p < 0.05$, one-way ANOVA). The mean detected volume for the flexion angle that was used to generate the wear pool (30°) was only 50% of the true wear volume.

In all cases, the volume of falsely detected wear was small (< 15 mm^3) compared to the true wear volume (230 mm^3). At a flexion angle of 30°, no falsely detected wear was found for all RSA image pairs.

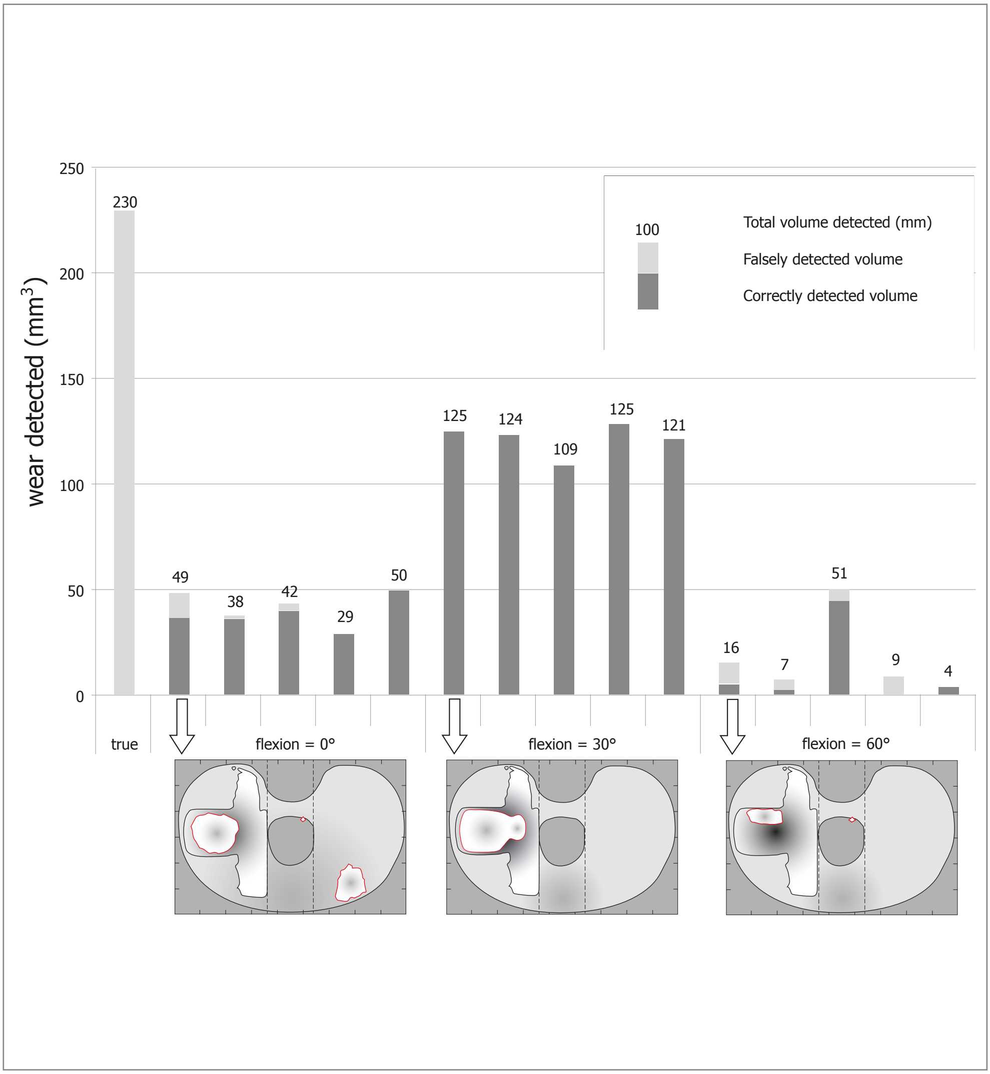


Figure 6-22. The wear volumes detected for each of the fifteen RSA image pairs, compared to the true wear volume (bar on the left). The images on the bottom of the figure are examples of the wear pools per flexion angle. The dark grey overlay indicates the true wear pool and the grey overlay on top indicates the detected wear pool.

The standard deviation at 0°, 30° and 60° of knee flexion were 8 mm^3, 7 mm^3 and 18 mm^3, respectively. The corresponding 95% prediction intervals, which are a measure for influence of variation in the femoral positioning, ranged between 12% and 33% of the volume of the true wear pool.

Table 6-12. The volumes detected in mm3 per flexion angle (N=5). The means and standard deviations (SD) for the total detected volume, correctly detected, falsely detected volume and missed volume are shown. The 95% prediction intervals give the expected variation in practice and are calculated as 4 x SD of the total detected wear volume.

Volumes in mm³	**Flexion = 0°**		**Flexion = 30°**		**Flexion = 60°**	
	mean	**SD**	**mean**	**SD**	**mean**	**SD**
Total detected	42	8	122	7	17	19
Correctly detected	38	7	122	7	11	20
Falsely detected	3	5	0	0	7	3
Missed	192	7	109	7	220	20
Prediction interval – size (relative to total volume) in mm³						
	Flexion = 0°		Flexion = 30°		Flexion = 60°	
	32	(14%)	28	(12%)	76	(33%)

6-4 Discussion

We investigated the influence of the pose estimation error, insert thickness deviation and variation of the tibiofemoral contact location on the accuracy and precision of the volumetric wear measurement. We found that pose estimation was the largest source of variation, producing a variation between 111 to 283 mm^3 (95% prediction interval). This equaled 100% up to 200% for smaller wear pools relative to the detected wear volume.

The 95% prediction interval due to insert thickness deviation was between 74 to 174 mm^3, which was smaller than the effect of pose estimation error. An important difference between these error sources is that pose estimation error influences each measurement randomly whereas the error due to insert thickness deviation is constant per patient. So in relative measurements to determine the wear progression, the error due to insert thickness variation is negligible.

Concerning variation due to femoral positioning, the repeated measurements in the phantom experiment (n = 5) showed an average SD of 10 mm^3, which is equivalent to a PI of 40 mm^3, i.e. 17% of the true wear pool volume.

The measurement accuracy in the phantom experiment was very limited as even in the best case only 56% of the wear pool volume was detected (129 of 230 mm^3, 30° of flexion). Moreover, for some of the inserts we were unable to detect any wear (Appendix 1). For some cases the low accuracy may be caused by a large distance between the tibiofemoral contact location and the center of the wear pool. A positive finding was that the falsely detected volumes were low (< 15 mm^3), resulting in a low risk of overestimating the wear pool.

A limitation of our study is that the validation is based on phantom and in silico data only, whereas the ideal validation would be based on RSA data from patients shortly before insert revision, ensuring that both the shape of the wear pool in the retrieved inlay and the femoral-insert contact location in the pre-op RSA image are representative. As such data was not available a phantom experiment was utilized in which both the shapes of the wear pool and the freedom of the tibiofemoralcontact location could be controlled to mimic clinical conditions. It is likely that the underestimation of the wear pool size and limited reliability found in this study are representative for clinical practice, as the created wear pools were a reasonable reproduction of abrasive wear.

Gill et al. suggested superimposing assessments with volumetric wear measurements at different flexion angles to get a better detection of the wear pool [48]. The findings in our phantom study confirm that superimposing assessments can be beneficial, as large differences were found in the detected wear volume among the flexion angles. However, our simulation study also showed that a single assessment already has a variation of 111 to 283 mm^3. When several (almost) disjoint wear pools detected in alternate flexion angles are superimposed, the total variation will further increase. In practice, we expect a tradeoff between the accuracy (underestimation) and the precision of the measurement. Repeated measurements for each flexion angle could be used to improve the precision, but then the required number of RSA acquisitions quickly becomes impractical.

In summary, the accuracy of the volumetric wear measurement was limited, as at most 56% of the true wear volume was detected. In addition, the precision of the measurement was low, mainly caused by the pose estimation. The prediction interval of the volumetric wear measurement is at least as large as the measurement outcome itself. This is an important limitation to the applicability of the volumetric wear measurement in clinical practice and further clinical validation is required.

References

[16] Naudie, D.D.R., et al., *Wear and osteolysis around total knee arthroplasty.* The Journal of the American Academy of Orthopaedic Surgeons, 2007. **15**: p. 53-64.

[20] Sundfeldt, M., et al., *Aseptic loosening, not only a question of wear: a review of different theories.* Acta Orthopaedica, 2006. **77**: p. 177-97.

[22] Sharkey, P.F., et al., *Why are total knee arthroplasties failing today?* Clinical Orthopaedics and Related Research, 2002. **404**: p. 7-13.

[27] Miller, T.T., *Imaging of knee arthroplasty.* European Journal of Radiology, 2005. **54**: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, *Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses.* The Journal of Arthroplasty, 2003. **18**: p. 860-866.

[30] Sanzén, L., et al., *Radiographic wear assessment in a total knee prosthesis.* The Journal of Arthroplasty, 1996. **11**: p. 738-42.

[32] Kaptein, B.L., et al., *A new model-based RSA method validated using CAD models and models from reversed engineering.* Journal of Biomechanics, 2003. **36**: p. 873-882.

[33] Kaptein, B.L., et al., *Clinical validation of model-based RSA for a total knee prosthesis.* Clinical Orthopaedics and Related Research, 2007. **464**: p. 205-9.

[40] Duryea, J., et al., *Fully automated software to monitor wear in prosthetic knees using fluoroscopic images.* European Radiology, 2001. **11**: p. 2184-7.

[48] Gill, H.S., et al., *In vivo measurement of volumetric wear of a total knee replacement.* The Knee, 2006. **13**: p. 312-7.

[76] van IJsseldijk, E.A., et al., *A model-based approach to measure the minimum joint space width of total knee replacements in standard radiographs.* Journal

of Biomechanics, 2012. **45**: p. 2171-5.

[84] Psychoyios, V., et al., *Wear of congruent meniscal bearings in unicompartmental knee arthroplasty.* The Journal of Bone & Joint Surgery. British Volume, 1998. **80**: p. 976-982.

[85] Edwards, S.a., et al., *Analysis of polyethylene thickness of tibial components in total knee replacement.* The Journal of Bone & Joint Surgery. American Volume, 2002. **84-A**: p. 369-71.

7

3D measurement of joint space narrowing in the knee from stereo radiographs using statistical shape models

Three dimensional measurement of joint space narrowing in the knee from stereo radiographs using statistical shape models

E.A. van IJsseldijk[a], E.R. Valstar[a,b], B.C. Stoel[c], R.G.H.H. Nelissen[a], N. Baka[c], R. van 't Klooster[c,1], B.L. Kaptein[a]

[a] *Biomechanics and Imaging Group, Department of Orthopaedics, Leiden University Medical Center, The Netherlands*

[b] *Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, The Netherlands*

[c] *Department of Radiology, Division of Image Processing, Leiden University Medical Center, The Netherlands*

As accepted by the Journal of Bone and Joint Research, April 2016

1 Currently working with Quantib BV, Westblaak 106, 3012 KM Rotterdam, The Netherlands

Abstract

Introduction

An important measure for the diagnosis and monitoring of osteoarthritis of the knee is joint space narrowing (JSN), which is assessed from plain radiographs by measuring longitudinal changes in the minimum joint space width (mJSW). Conventional 2D mJSW measurements require alignment of the X-ray beam with the surface of the medial tibial plateau. We propose a newly developed mJSW measurement technique from stereo radiographs using 3D statistical shape models (SSM) of the tibia and femur and evaluate its sensitivity to changes in the mJSW and its robustness to variations in patient positioning and bone geometry.

Method

A validation study was performed using cadaver specimen for which the actual mJSW could be varied using a micromanipulator. For comparison purposes, the mJSW was also assessed from plain radiographs using the conventional 2D measurement method. To study the influence of SSM model accuracy, the 3D mJSW measurement was repeated with bone models obtained from CT scans.

Results

The SSM-based measurement method was more robust than the conventional 2D method, showing that the 3D reconstruction indeed reduces the influence of patient positioning. Both methods showed comparable sensitivity to changes in mJSW. The CT-based measurement was more accurate than the SSM-based measurement, (smallest detectable differences 0.55 vs. 0. 82 mm respectively) indicating that the modelling error of the SSM is probably an important contributor to SSM measurement accuracy.

Conclusion

In conclusion, the proposed measurement method is not a substitute for the conventional 2D measurement as it is more complicated to conduct and its improvements on measurement accuracy are marginal. However, further improvement of the model accuracy and optimization technique can be obtained and will stimulate applicability.

7-1 Introduction

Osteoarthritis (OA) of the knee imposes a major health care burden with a reported prevalence of more than 18% in the 65 to 74 year age group in the European Union [86]. OA is associated with cartilage degeneration and loss, joint inflammation, and swelling of the joint. Patients experience pain, stiffness and limited mobility [87].

OA progression is most frequently evaluated using plain radiographs for their low costs and availability. A variety of features are used to assess the stage of OA, such as the appearance of osteophytes and subchondral sclerosis [74, 88]. Cartilage loss associated with OA is estimated by detecting joint space narrowing (JSN). This is measured based on longitudinal changes in the minimum joint space width (mJSW), i.e. the shortest visible distance between the femoral condyle and the tibial plateau.

A limitation of plain radiographs is that measurements are conducted in projection views that are prone to parallax effects. As a result, alignment of the X-ray beam with the surface of the medial tibial plateau (MTP) is crucial in order to obtain a reliable reading of the joint space [89, 90]. Standardization protocols have been developed to optimize alignment, such as the fixed-flexion (FF) view, the metatarsophalangeal view and the modified Lyon Schuss view [91, 92]. An alternative measurement approach could be to reconstruct the three-dimensional (3D) bone geometry around the knee joint from planar images. This reconstruction has the advantage that geometric measurements such as the mJSW are invariant to the projection angle. This reduces the influence of variation in patient positioning or bone geometry and improves the accuracy and precision of the measurement.

We therefore developed a technique, in which this 3D reconstruction is created from 3D shape models of the tibia and femur and 2D/3D matching in Roentgen Stereophotogrammetric Analysis (RSA) [32, 34]. Afterwards, the mJSW is measured with a similar technique used for the measurement of polyethylene wear in total knee prostheses [72]. A particular challenge of this 3D reconstruction is that patient-specific 3D models of the tibia and femur are not readily available. To solve this, 3D statistical shape models (SSMs) of the tibia and femur were developed. An SSM is a deformable model that incorporates shape variations of an object class from a training set of examples. These models could be used to produce accurate reconstructions of 3D patient-specific bone shapes based on 2D image information [93].

In this study, the feasibility of this newly developed mJSW measurement technique was investigated. A validation study was performed using cadaver specimen for which the actual mJSW could be varied using a micromanipulator. For comparison, the mJSW was also measured in conventional plain radiographs with optimized medial tibial plateau alignment, using an image-based semi-automatic measurement technique [94]. To study the influence of model accuracy, the 3D mJSW measurement was repeated with bone models obtained from CT scans and with the SSMs.

7-2 Materials and Methods

7-2-1 Data

Five human, cadaveric legs with no visible pathology were selected from the Department of Anatomy of the Leiden University Medical Center (Table 7-13). All ligaments and soft tissues including cartilage were dissected so that only the naked tibia and femur remained.

Table 7-13. Main characteristics of the cadaveric specimen. Age is expressed in years.

index	Gender	Age	Leg Side
1	Female	91	Right
2	Male	98	Right
3	Female	63	Right
4	Female	93	Left
5	Male	84	Right

7-2-2 Models

3D CT models from cadaver bones

3D surface models of the cadaveric bones were created from helical CT scans (Toshiba Aquilion 64, Toshiba Medical Systems Ltd., Tokyo, Japan). The bones were arranged in such a way that their long axes were aligned parallel to the CT table. The bones were separated using foam padding in order to simplify the digital delineation of the bones. The scans were obtained at 120 kV and 130 mA with a slice thickness of 1.0 mm and a pitch of 0.8 mm per revolution. The scans had a resolution of 512 by 512 by 641 voxels with a voxel size of 0.78 by 0.78 by 0.8 mm.

Image segmentation was employed using Amira software (FEI Visualization Science Group, Bordeaux, France). A voxel mask was created to separate the bones from the background in the CT images using a threshold-based approach. The mask was converted into a triangulated surface model using a marching cube algorithm [95]. The average triangle edge length of the models was 1.7 mm.

Statistical Shape models

An SSM is a deformable model of shape that learns the mean shape and likely shape variations of an object class based on a training set. It can generate new shapes using the formula, $x = \bar{x} + \Phi b_x$, where $\bar{x}$ is the mean shape of the training set, Φ is the set of eigenvectors (modes of variation) that is based on the covariance matrix of the training set and b is the set of shape parameters, one for each eigenvector. Thus, b_x stands for the set of parameter values corresponding to the generated shape x [96].

In this work two SSMs were used to model the distal femoral and proximal tibial bones, truncated to the region near the knee joint (each approximately 12cm in length). The two models originate from a previous study where they are described in detail [97]. The training sets consisted of 62 polygonal surface models that were created from CT data using a level-set segmentation. Correspondence in the training sets was achieved using a non-rigid registration with the Elastix software [98]. Note that the five cadaver bones from this study are not included in the training set.

The eigenvector sets of the SSMs were truncated so that only those modes remained that describe 95% of the eigenvalue sum. For both models 33 modes of variation remained. For each mode j, the corresponding shape parameter b_j was allowed to vary between ±3 times the standard deviation (SD) of the corresponding eigenvalue ($-3\ SD_j \leq b_j \leq 3\ SD_j$) when generating new shapes.

To test the goodness of fit, the models were fitted to each of the 3D surface models of the cadaveric bones using 3D/3D matching and the root mean square point-to-surface distance was computed. The root mean square point-to-surface distances ranged between 0.49 and 0.74 mm, which indicates that results are similar to earlier studies using SSMs [97].

7-2-3 mJSW measurement methods

In this section, the mJSW measurement methods are described for the SSM-based measurement, the conventional 2D measurement and the CT based measurement (Figure 7-23).

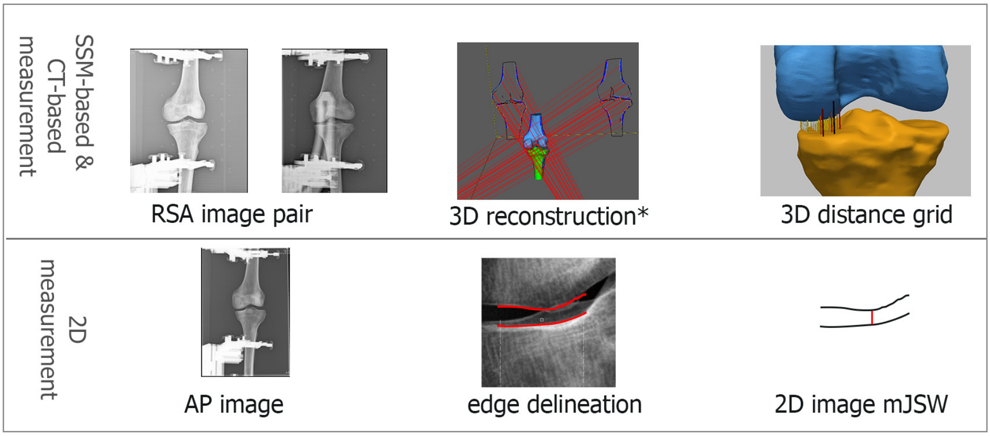


Figure 7-23. Schematic view of the intermediate steps of the mJSW measurement methods. The steps start with the original images and end with the feature that is used to compute the mJSW. The 3D reconstruction step also includes the model optimization, which differs between the SSM-based and CT-based measurements.

SSM-based measurement

The SSM-based measurement is conducted using RSA image pairs. In essence, a 3D reconstruction of the femur and tibia is created, in which the mJSW is measured.

First, image calibration and edge delineation are done using a standard analysis in Model-based RSA software (v4.0, LUMC, Leiden, Netherlands). In this analysis, candidate edges are detected with a canny-edge-detection algorithm and a selection is made semi-automatically. To avoid correspondence problems, only those edges were selected that a) represented the outer object contours and b) belonged to the region that the SSM could represent (i.e. the distal part of the femur and proximal part of the tibia).

The next step is to optimize the shape parameters as well as the pose parameters of the tibiae and femora. This optimization step is done in MATLAB (R2011a) using a 2D/3D matching algorithm. Validation of this algorithm in previous work found a root-mean-square error of 1.86±0.29 mm for the femoral model [97]. The tibial bone was not included in this validation experiment.

Last, the mJSW is computed as the minimum distance between the tibia and the femur model. This distance is measured in the direction perpendicular to a 0.2 mm by 0.2 mm measurement grid residing in the transverse plane beneath the medial condyle (Figure 7-24). The construction of the measurement grid and the coordinate system of the tibia is based on three landmark regions manually defined on the tibial SSM model. These regions transform with the shape optimization, so that this procedure is required only once.

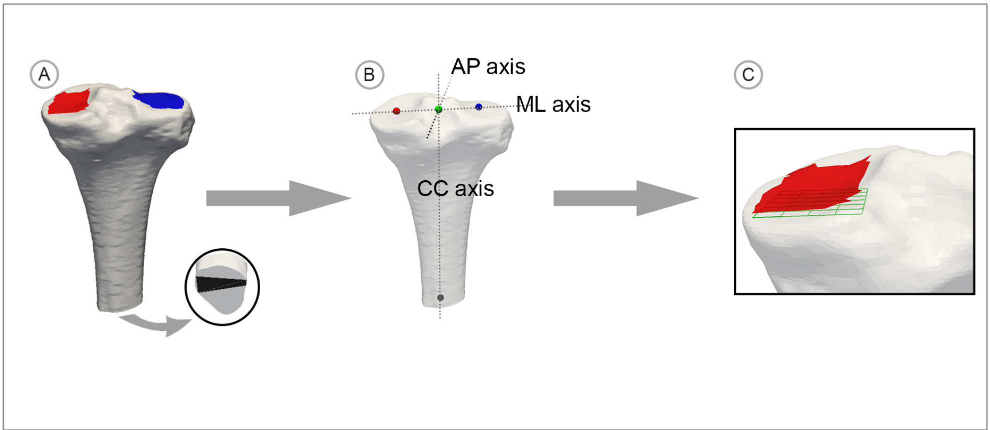


Figure 7-24. illustration of the grid construction process. A) Three tibia surfaces areas are selected by the used. B) The geometric means of these locations are used to define the coordinate system. C) The measurement grid is constructed beneath the medial condylar surface area aligned with the coordinate system.

2D measurement

The 2D measurement was performed with an automatic technique which has been validated for mJSW measurements in hand radiographs (freely available at www.lkeb.nl). The smallest detectable difference (SDD = 1.96 x SD) ranged between 0.05 mm and 0.354 mm depending on the joint shape [94]. This technique was adopted for the current measurement in terms of image contrast and joint size. Hereto, the proximal (femoral) and distal (tibial) margins of the medial knee joint are delineated using a semi-automatically algorithm specialized for these structures [94]. The user selects the center point of the medial tibial plateau in the image and the algorithm returns the edges of the margins in a 20 mm range. Optionally, the user can provide additional guiding points to correct these edges manually. The shortest perpendicular distance within the interval of delineation divided by image magnification was stored as the mJSW.

CT-based measurement

A CT-based measurement was used to study the influence of model accuracy. This measurement used models based on CT-scans instead of the SSM models. The calibration and edge selection for the CT-based measurement are similar to the SSM-based measurement. In the 2D/3D matching step however, the pose parameters (position, orientation, isotropic scale) of the CT models of the tibiae and femora are optimized using the default 2D/3D matching algorithm in Model-based RSA software.

7-2-4 Experiments

A validation experiment was done using a set-up in which the actual medial mJSW of the cadavers could be controlled with a micro manipulator (Figure 7-25) as part of a positioning device (accuracy 0.01 mm). This set-up was used to acquire both plain radiographs and RSA images under equal, controlled circumstances.

The plain radiographs were acquired with an X-ray imaging system at the Leiden University Medical Center (CXDI-series, 169dpi, 12BPP, Canon, New York, USA). A standing anterior-posterior (AP) view was used with a focus-film distance of 1.2 meters. The image magnification factor was 110%, based on measurements of the bone to detector distances. For the RSA images a mobile X-ray system with the same

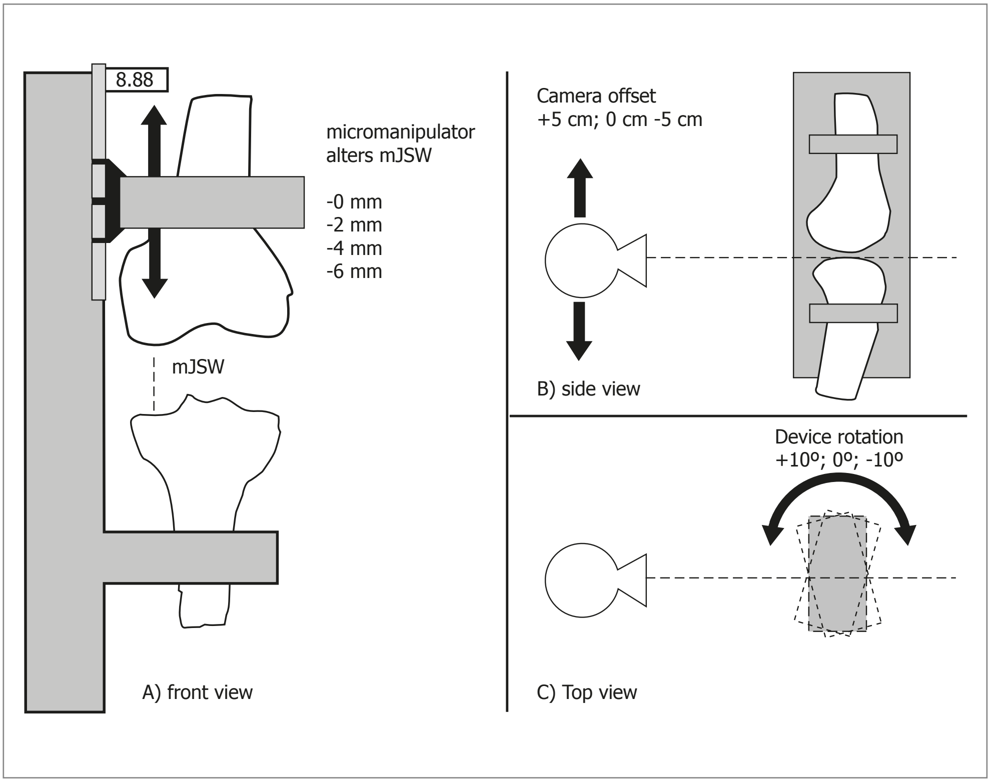


Figure 7-25. A) Schematic view of the positioning device and manipulation of the actual mJSW. B) Effect of manipulating the X-ray tube offset parameter. C) Effect of manipulating the rotation of the positioning device.

device qualifications was added. The detectors were placed in a carbon calibration box (LUMC, Leiden, The Netherlands). The X-ray sources were positioned at 1.5 meters from the detectors and the angle between the X-ray beams was approximately 40°. For both imaging modes the positioning device was placed as close to the detectors as possible.

For each cadaver, first a plain radiograph was acquired with the actual mJSW at 0 mm, i.e. in which there was contact between the medial femoral condyle and the tibial plateau. For this acquisition, the medial tibial plateau was aligned with the X-ray beam. This was achieved by optimizing the positioning of the tibia in the phantom and by adjusting the height of the X-ray tube until the center of the beam (laser guidance) just skimmed the edges of the plateau. Since alignment was optimized for the medial plateau, the lateral mJSW was not measured in this validation study.

After the first acquisition at 0 mm, the actual mJSW was increased to 2 mm, 4 mm and 6 mm. These values are representative for diseased and healthy adult knees. For each of these distances, 3 exposures were made with varying X-ray tube heights (-5 cm, 0 cm, +5 cm) and 3 exposures were made with varying rotations of the cadaver (-10°, 0°, +10°). Note that when one parameter was varied, the other parameter was in neutral position and the exposure with zero tube height and zero rotation was made twice. In total, 19 exposures were made per cadaver. A schematic with the function of these parameters is shown in Figure 7-25. The range of the parameters was considered representative for actual variations in patient positioning during follow-up studies.

The above procedure was repeated acquiring RSA image pairs for each cadaver bone. The mJSW was measured using the 2D measurement in plain radiographs and using the SSM-based and CT-based measurements for the RSA image pairs, resulting in 285 measurements in total.

7-2-5 Statistical analysis

From the experiment data, the relative measurement errors were computed as the measured mJSW minus the actual mJSW. To analyze the robustness of the measurements against the variations in position applied in the experiment, the measurement errors per method are shown in a boxplot. Significant differences between the dispersion are tested with Levene's test.

The sensitivity was evaluated based on the data with mJSW variation only. Measurements with a mJSW of 0 mm and with any tube offset or rotation were excluded, (N = 6 measurements per cadaver). Standard deviations (SD) and the smallest detectable differences (SDD = 1.96xSD) were computed. The SDD is a relevant outcome for OA research, representing the minimum JSN that could be detected [88, 99]. Between-cadaver differences were analyzed with a univariate linear model with the shape index as random factor. Last, data trends were analysed by plotting the measurement errors against the actual mJSW.

7-3 Results

In the robustness analysis, the measurement errors showed a large difference in dispersion between the measurement methods (Figure 7-26). Generally, the smallest dispersion was found for the CT-based measurement, next for the SSM-based measurement and last for the 2D measurement. These differences were statistically significant for cadaver 2 to 5 (Levene's test, $p < 0.01$).

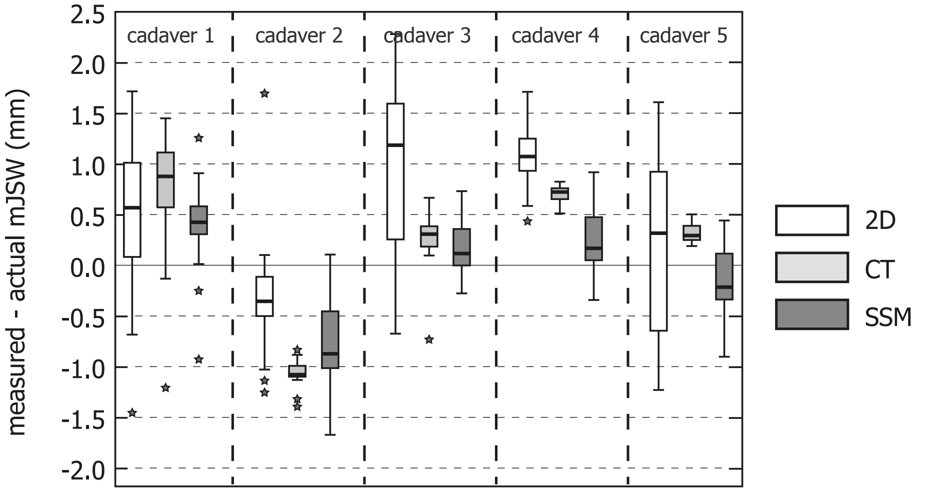


Figure 7-26. Boxplots presenting the difference between the actual mJSW and measured mJSW in the validation experiment for each method and cadaver shape (N = 19 for each boxplot). The horizontal bar indicates the median difference. The whiskers are set at 1.5 times the interquartile range.

No significant differences or trends were found between the measurement error versus the actual mJSW (Figure 7-27). The results in Table 7-14 show that the SDDs differed significantly between the cadavers for all measurement methods (ANOVA, $p < 0.01$). More specifically, the SDD of cadaver 1 was relatively high for all measurement methods. The last table column shows the SDDs corrected for between-cadaver effects with the univariate linear model. This shows the corrected SDD is smallest for the CT-based measurement method, followed by the 2D measurement and the SSM-based measurement method respectively.

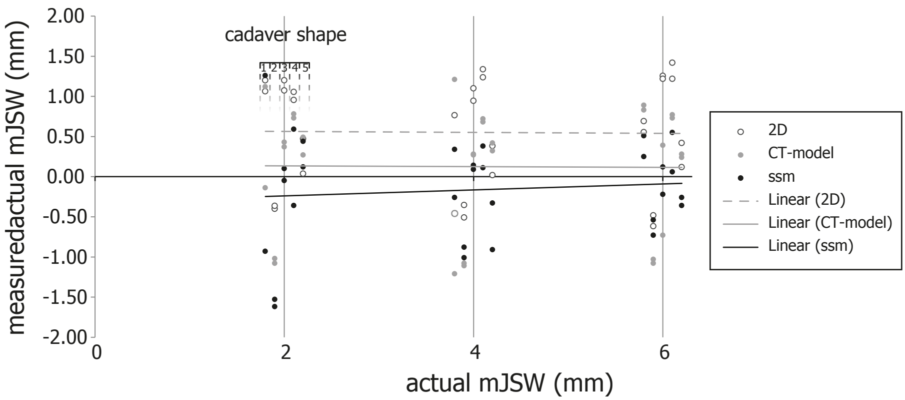


Figure 7-27. The measurement errors as a function of the actual mJSW (tube offset 0 cm, rotation 0 degree) together with linear trendlines. To improve the readability of the plot, the dots have a slight horizontal offset based on the index of the cadaver bones as illustrated at mJSW = 2 mm.

Table 7-14. The standard deviations and smallest detectable differences of the three measurement methods (with tube offset 0 cm, and rotation 0 degree). In the columns, values are first shown per cadaver and then for the whole dataset.

		Single cadavers (1 to 5)					**Whole dataset***
		N = 6	**N = 6**	**N = 6**	**N = 6**	**N = 6**	**N = 30**
Standard deviation	**2D**	0.59	0.10	0.12	0.17	0.21	0.36
(SD)	**3D – CT**	0.95	0.03	0.44	0.04	0.10	0.28
in mm	**3D – SSM**	0.74	0.44	0.14	0.36	0.46	0.42
Smallest detectible	**2D**	1.15	0.20	0.23	0.34	0.40	0.70
difference (SDD)	**3D – CT**	1.85	0.07	0.87	0.07	0.19	0.55
in mm	**3D – SSM**	1.45	0.85	0.27	0.70	0.90	0.82

* based on the least square error in the univariate analysis

7-4 Discussion

The purpose of this article was to investigate the application of a SSM reconstruction of the knee to conduct mJSW measurements and assess the feasibility of this measurement method. In a validation study, its sensitivity to changes in the mJSW was evaluated as well as its robustness to variations in knee positioning and bone geometry. For comparison, the mJSW was also measured in conventional plain radiographs. The measurement was repeated with bone models obtained from CT scans to study the influence of model accuracy. In comparison with the conventional 2D mJSW measurement from plain radiographs, the method is more robust (Figure 7-26) with a similar sensitivity over the whole dataset (Table 7-2). Thus, 3D reconstruction reduces the influence of knee positioning as expected. However, we could not establish an improvement in sensitivity, since the SDD of the SSM-based measurement is higher than that of the conventional measurement and CT-based measurement (SDD = 0.82 mm, 0.7 mm and 0.55 mm respectively). The error in the SSM-based measurement can originate from different sources: image calibration error, edge detection error, fitting error (i.e. not finding the global minimum solution) and modelling error. The comparison of results of the SSM-based measurement and the CT-based measurement shows that the CT-based measurement results were more accurate than those of the SSM-based measurement. This indicates that modelling error is probably an important contributor. The modelling error could be reduced by increasing the training set of the shape models. Another option is to improve the 2D/3D fitting and optimization. For example, edge detection can be inaccurate or incomplete when parts of the femoral and tibial silhouettes overlap. This could be solved by searching for better edge candidates in the neighborhood of the SSM silhouette during optimization. In addition, the edge orientation can be used to discriminate between the femoral and tibial edges, which is a technique that already has been studied [36]. Moreover, in clinical practice follow-up images are available. These can be exploited to limit the search space in which the optimization is performed.

The validation experiment showed that a 3D reconstruction improves the robustness of the measurements against variations in patient positioning, which was simulated using different tube offsets and rotation angles of the knee. Although this comparison is useful, the robustness found for the 2D measurement cannot be extrapolated directly to clinical practice as measurement protocols are often employed reducing variability. Also, images that violate certain specifications (such as a high inter-margin distance) are retaken, further reducing variations in viewing angles. These protocols reduce variations in patient positioning and viewing angles by reducing inaccuracies in mJSW measurements in clinical practice.

A curious finding was that cadaver 1 showed relatively high measurement errors for all methods. This could be caused by bone abnormalities in the medial joint shape. As can be seen in Figure 7-28, a bulge was present in the femoral bone as well as a large inter-margin difference in the tibial plateau. Although the CT-based measurement does incorporate this bulge, results still show high measurement errors, indicating that other factors influence the measurement results.

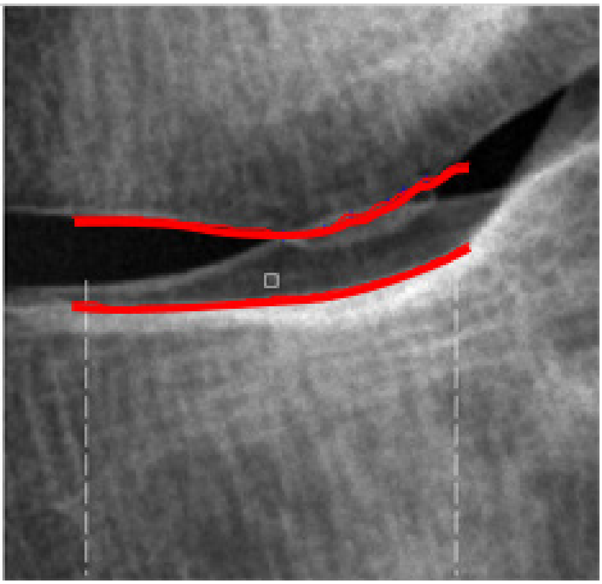


Figure 7-28. Screenshot of the contour delineation for the 2D measurement in one of the examinations for cadaver 1.

The reproducibility of the 2D mJSW measurement has been evaluated in several other studies. Dupuis et al. found an SD of 0.08 mm to 0.11 mm in a cadaver study [100], which is a remarkably high precision. Conrozier et al. reported an SD of 0.14 mm for the reproducibility when fluoroscopy-assisted radiographs were used [90]. Except for the first cadaver specimen, the results in our study are comparable with an SD ranging between 0.10 and 0.20 mm.

Only cadaver knees without signs of OA were used in this validation study, because it was designed as a proof of concept of the measurement method. For patients with OA, modelling the femoral and tibial bones will be more challenging, because of abnormal shapes and the formation of osteophytes. The optimization of shape and pose parameters in the SSM-based measurement can be adjusted for such aberrations. For example, semi-automatic or automatic detection of the corresponding regions can be introduced, followed by the assignment of different weights to these regions in the 2D/3D matching algorithm.

More sophisticated imaging techniques such as MRI and CT are considered as a more reliable alternative than planar radiographs for the estimation of cartilage loss [101, 102]. However, these methods are more costly, more time-consuming and require experience and special equipment. Given that the modelling error can be further improved, the SSMs can provide a good alternative. Moreover, SSMs can provide quantitative information on the bone morphology. This has proven its value in the identification of risks and in the diagnosis of skeletal diseases [93, 103]. For example, the risks for hip fractures, the progression of osteoarthritis of the hip and the need for total hip replacement can be estimated by analysing the shape of the femur using a SSM model [104-106]. Likewise, a SSM-based reconstruction of the knee can be used to combine shape analyses and geometric measurements such as the mJSW, which can be valuable for OA-related research [107].

This study focused only on the validation of the mJSW measurement, but the 3D reconstruction has other contributions as well. For example, the 3D location of the mJSW could be determined and correlations between progression of joint space narrowing and (changes in) the 3D bone geometry can be studied. Also, alternative metrics such as the median or mean joint space distance could be investigated.

These metrics are probably less susceptible to noise or outliers than the mJSW, but often do require a standardized definition of the tibial margin based and deviate from the current definition of JSW.

In conclusion, the proposed measurement method is not a substitute for the conventional 2D measurement. The marginal improvement in measurement accuracy does not outweigh the increase in measurement complexity. However, further improvement of the model accuracy and optimization technique can be obtained. Combined with the promising options for applications using quantitative information on the bone morphology, SSM based 3D reconstructions of natural knees are interesting for further development.

References

[32] Kaptein, B.L., et al., *A new model-based RSA method validated using CAD models and models from reversed engineering.* Journal of Biomechanics, 2003. **36**: p. 873-882.

[34] Ryd, L., et al., *Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses.* The Journal of Bone & Joint Surgery. British Volume, 1995. **77**: p. 377-83.

[36] Baka, N., et al., *2D-3D shape reconstruction of the distal femur from stereo X-Ray imaging using statistical shape models.* Medical Image Analysis, 2011. **15**(6): p. 840-850.

[72] van IJsseldijk, E.A., et al., *The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA.* Journal of Biomechanics, 2011. **44**: p. 2724-7.

[74] Buckland-Wright, C., *Which radiographic techniques should we use for research and clinical practice?* Best Practice & Research in Clinical Rheumatology, 2006. **20**: p. 39-55.

[86] Schaap, L.A., et al., *European Project on OSteoArthritis (EPOSA): methodological challenges in harmonization of existing data from five European population-based cohorts on aging.* BMC musculoskeletal disorders, 2011. **12**(1): p. 272.

[87] Shapiro, J.R., et al., *Osteogenesis imperfecta: A translational approach to*

*brittle bone disease*2013: Academic Press.

[88] Guermazi, A., et al., *Imaging in osteoarthritis.* Rheumatic diseases clinics of North America, 2008. **34**: p. 645-87.

[89] Mazzuca, S.A., et al., *Effect of alignment of the medial tibial plateau and X-ray beam on apparent progression of osteoarthritis in the standing anteroposterior knee radiograph.* Arthritis & Rheumatism, 2001. **44**: p. 1786-1794.

[90] Conrozier, T., et al., *Influence of the quality of tibial plateau alignment on the reproducibility of computer joint space measurement from Lyon schuss radiographic views of the knee in patients with knee osteoarthritis.* Osteoarthritis and Cartilage, 2004. **12**: p. 765-70.

[91] Mazzuca, S.A., et al., *Performance of a non-fluoroscopically assisted substitute for the Lyon schuss knee radiograph: quality and reproducibility of positioning and sensitivity to joint space narrowing in osteoarthritic knees.* Osteoarthritis and Cartilage, 2008. **16**: p. 1555-9.

[92] Buckland-Wright, J., et al., *Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views.* The Journal of Rheumatology, 1999. **26**(12): p. 2664-2674.

[93] Sarkalkan, N., H. Weinans, and A.A. Zadpoor, *Statistical shape and appearance models of bones.* Bone, 2014. **60**: p. 129-40.

[94] Huétink, K., et al., *Automatic radiographic quantification of hand osteoarthritis; accuracy and sensitivity to change in joint space width in a phantom and cadaver study.* Skeletal Radiology, 2012. **41**: p. 41-9.

[95] Lorensen, W.E. and H.E. Cline. *Marching cubes: A high resolution 3D surface construction algorithm*. in *ACM Siggraph Computer Graphics*. 1987. ACM.

[96] Cootes, T.F., et al., *Use of active shape models for locating structures in medical images.* Image and vision computing, 1994. **12**(6): p. 355-365.

[97] Baka, N., et al., *Evaluation of automated statistical shape model based knee kinematics from biplane fluoroscopy.* Journal of Biomechanics, 2014. **47**(1): p. 122-129.

[98] Klein, S., et al., *Elastix: a toolbox for intensity-based medical image registration.* Medical Imaging, IEEE Transactions on, 2010. **29**(1): p. 196-205.

[99] Bland, J.M. and D.G. Altman, *Statistics notes: measurement error.* Bmj, 1996. **313**(7059): p. 744.

[100] Dupuis, D.E., et al., *Precision and accuracy of joint space width measurements of the medial compartment of the knee using standardized MTP semi-flexed radiographs.* Osteoarthritis and Cartilage, 2003. **11**(10): p. 716-724.

[101] Eckstein, F., T. Mosher, and D. Hunter, *Imaging of knee osteoarthritis: data beyond the beauty.* Current Opinion in Rheumatology, 2007. **19**: p. 435-43.

[102] Wirth, W., et al., *Direct comparison of fixed flexion, radiography and MRI in knee osteoarthritis: responsiveness data from the Osteoarthritis Initiative.* Osteoarthritis and Cartilage, 2013. **21**(1): p. 117-125.

[103] Heimann, T. and H.-P. Meinzer, *Statistical shape models for 3D medical image segmentation: a review.* Medical Image Analysis, 2009. **13**: p. 543-63.

[104] Gregory, J.S., et al., *A method for assessment of the shape of the proximal femur and its relationship to osteoporotic hip fracture.* Osteoporosis International, 2004. **15**: p. 5-11.

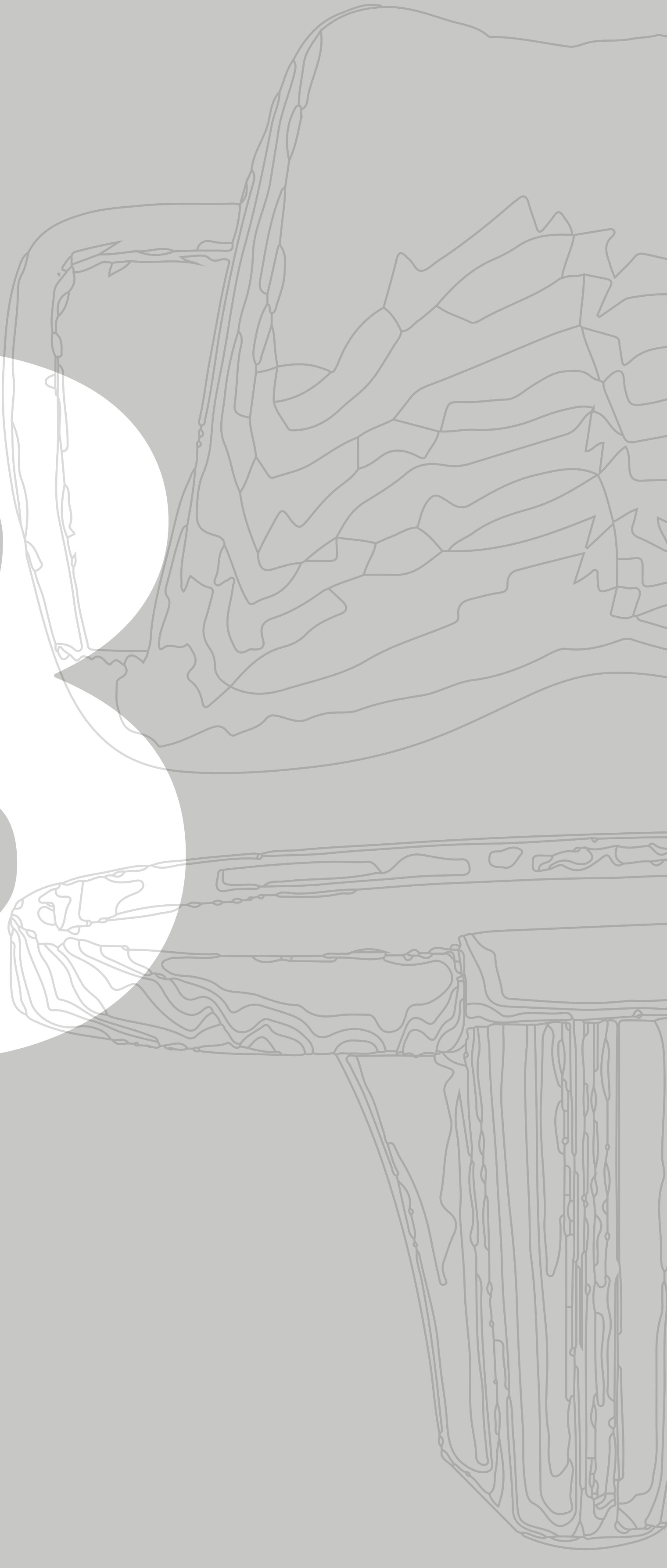
[105] Barr, R.J., et al., *Predicting OA progression to total hip replacement: can we do better than risk factors alone using active shape modelling as an imaging biomarker?* Rheumatology, 2012. **51**(3): p. 562-570.

[106] Lynch, J.A., et al., *The association of proximal femoral shape and incident radiographic hip OA in elderly women.* Osteoarthritis and Cartilage, 2009. **17**(10): p. 1313-1318.

[107] Felson, D.T. and T. Neogi, *Osteoarthritis: is it a disease of cartilage or of bone?* Arthritis & Rheumatism, 2004. **50**(2): p. 341-344.

8

Discussion



Discussion

8-1 General discussion

Total Knee Arthroplasty (TKA) is an effective treatment for end stage symptomatic osteoarthritis with good long-term outcomes and patient satisfaction [10, 11, 14, 108]. Survival analysis taking revision of the implant as end-point shows a mean survival at 10 years of 90%. The main cause for revision is loosening of the tibial or femoral components, which is related to wear of the polyethylene liner insert [9, 12, 13, 15, 20]. As the rate at which the polyethylene insert thickness decreases can predict failure [23], an accurate and precise method is required to assess the progression of this wear *in vivo*. This will not only support clinical decision making on when to exchange a liner before component loosening occurs but also enables the accurate comparison of wear resistance of different (new) prosthetic designs [25].

In current clinical practice and research, the progression of the polyethylene wear is measured in planar radiographs using minimum joint space width (mJSW) measurement as a surrogate measure for the remaining insert thickness [27]. Measurement errors of up to 2 mm are not exceptional and multiple follow-up visits are required to obtain a reliable estimation of the wear rate [28, 29, 31]. Model-based measurement methods based on matching a 3D model of the prosthesis on its 2D projection in the roentgen image are less prone to human errors or parallax errors resulting from the alignment of the radiographic beam. It is therefore expected that a better accuracy and precision can be achieved using model-based wear measurement techniques.

The primary aim of this work was to develop novel model-based mJSW measurement methods and validate the accuracy and precision of these methods using conventional measurement methods as a reference. Next to the validation of the model-based measurements this work assessed the reliability of the mJSW measurement as a surrogate for the actual insert thickness as this reliability can be influenced by numerous factors. In this study, the influence of the image set-up and patient positioning were assessed.

mJSW measurements for TKAs

Chapters 2 and 3 focus on *in vitro* validation of the developed model-based mJSW measurement for TKAs using either stereo-images acquired from RSA or anteroposterior (AP) plain radiographs: RSA is typically used to assess or compare new implant designs and is known for its high accuracy in 2D/3D pose reconstruction [33, 46]. AP radiographs are used in daily clinical practice (e.g. for the assessment of wear progression in TKA).

As expected, the highest precision and accuracy (0.2 mm and 0.1 mm respectively) can be achieved when stereo images are used (Chapter 2). The accuracy of model-based mJSW measurements with RSA was studied in a few existing studies where similar findings were reported [42, 43, 48]. This is the first study in which the influence of patient positioning on the accuracy of wear measurement was evaluated. Only anterior tilting showed a statistically significant effect (0.07 mm versus 0.02 mm accuracy for 0° versus 10° tilt). The use of reversed engineered (RE) models is preferred over CAD models because of the large increase in measurement accuracy and precision and the lower sensitivity for patient positioning and flexion angle. Recently, in a similar study no significant effects of variances in patient positioning were found, corresponding with our findings when using RE models [109].

For AP radiographs, the model-based technique has been applied using a standard imaging set-up without calibration object (Chapter 3). Compared to older studies that describe a similar model-based measurement method [41, 65], this has the advantage that the model-based technique can be immediately applied to any standard AP radiograph in clinical practice.

The model-based technique significantly improves the accuracy (0.2 versus 0.5 mm) and reproducibility (0.3 versus 1.0 mm) of mJSW measurements compared to the conventional measurements. These results indicate that a direct improvement of the mJSW measurement can be attained when applying the model-based measurement technique in clinical practice.

The mJSW measurement as a reliable estimator of the insert thickness
The mJSW is an indirect measurement of the insert thickness, which may not be reliable if the femoral component loses contact with the insert [110]. Furthermore, the insert thickness should be measured at the same location in successive follow-up images for reliable wear detection. This can be challenging as the femur generally performs a sliding and rolling motion over the articulating insert surface during flexion, yet only a single contact location is captured in an X-ray image. To detect wear, the contact location should coincide with the damaged area of the insert in the baseline radiographs as well as the successive follow-up radiographs. Capturing the damaged area can be challenging also as the location and size of this area can vary among patients [82].

The work presented in Chapters 4 and 5 improves our insight in the reliability of the mJSW as a wear indicator. Chapter 4 showed that the mJSW was larger in non-weight-bearing (NWB) than in weight-bearing (WB) images, with a mean difference of 0.28 mm and 0.20 mm for the medial and lateral condyle respectively. This difference can be explained by the possible loss of contact between the tibial and femoral components in NWB positions and differences in contact location between the WB and NWB positions. In NWB positions, the femoral contact location is more anterior with respect to the tibia due to gravity. As the insert is thicker at this location, this also explains the larger mJSW measured in NWB position. Patient positioning thus influences the mJSW measurement outcome and should be taken into account when interpreting these measurements.

In Chapter 5, the insert thickness of retrieved inserts was compared to model-based and conventional mJSW measurements in pre-operative weight-bearing radiographs. The model-based measurement has a higher accuracy than the conventional measurement, but the measurement precision was similar. This findings on precision differ from those in Chapter 3, where the model-based measurement was significantly more precise was (0.8 vs 0.2 mm standard deviation). In the study of Chapter 3 the thickness of flat acrylic blocks was measured in vitro whereas in Chapter 5 the measurement was used in vivo on actual inserts with a more complex articular surface. Therefore, the measurement precision in Chapter 5 can be influenced by differences between the insert measurement location and minimum insert location and by loss of tibiofemoral contact. In support of this, for five cases the medial mJSW measured by either technique was much larger (> 1 mm difference) than the actual insert thickness. We believe that the mJSW was measured accurately, but that the influences above resulted in a difference between measured and actual minimum insert thickness. The limited number of cases in this study did not allow for a detailed analysis of these effects, leaving the subject as an important topic for future work.

Insight in tibiofemoral location from model-based mJSW measurement

A major advantage of model-based mJSW measurement methods over the conventional method is the possibility to deduct the tibiofemoral contact location, i.e. the projection of the lowest point of each femoral condyle to the transverse plane of the tibial plateau. The mJSW measurement itself is conducted at this contact location. This information can be related to differences in the tibiofemoral contact location between subsequent measurements and can therefore be useful in clinical practice, e.g. to quantify the repeatability in successive follow-ups.

The accuracy and precision of this mJSW location measurement has not been assessed in this work. However, the tibial surface areas at which the mJSW locations were measured in our studies correspond to kinematics descriptions as well as retrieval studies describing insert surface damage patterns [79, 111, 112]. Moreover, in Chapter 5 we have shown that the mJSW location had a good correspondence with the location of the minimum insert thickness. In future work, the precision of the location measurement could be determined based on double examinations (test-retest image acquisition) with weight-bearing images.

Volumetric wear measurement for TKAs

Model-based reconstruction techniques can also be used to estimate the volume of wear debris as was already theorized by Gill et al. [48, 49]. This could be applied to in vivo performance testing as part of prospective evaluation of new implant designs in premarket release study, after in vitro wear simulator studies have been done. We developed such a volumetric wear measurement for TKAs and studied its accuracy using artificially worn liners. Measurements at different flexion angles (0, 30 and 45 degrees) were performed to investigate the influence of tibiofemoral contact and whether these measurements provide complementary information (Chapter 6). We found that the accuracy of this volumetric measurement is currently limited. Given the absolute outcomes of volumetric wear measurements, the influence of model positioning error is larger than for linear wear measurements. For example, bias in the pose estimation of the 3D models has a larger influence than in relative (baseline vs follow-up) linear wear measurements where this bias is cancelled out. In addition, the volumetric wear measurement also relies on the accuracy of the 3D insert model. The use of generic (non-prosthesis specific) 3D models results in a limited accuracy due to differences in the actual insert thickness from tolerance in the manufacturing process. In a similar experiment using a physiological phantom and inserts from retrievals, only half the volume of the total wear volume could be detected [113]. Given the current limitations in measurement accuracy, obtaining reliable volumetric wear measurements with this technique is not yet possible and linear wear measurements should be used instead.

Model-based mJSW measurement for native knees

The mJSW measurement is also used in radiographs of native knees to assess the progression of osteoarthritis [2, 114]. However, false readings may occur if the tibial plateau is skewed with respect to the X-ray beam [2, 115]. Moreover, cartilage defects are best detected when the images are acquired in a weight-bearing set-up and during in a flexion position of the knee. Due to this, the general opinion in literature is that this measurement lacks sensitivity and different approaches are advised such as measuring cartilage thickness on MRI or the use of a (fluoroscopy) guided imaging protocol to standardize patient positioning improving reproducibility [2, 35, 88, 100, 102, 115, 116].

In Chapter 7, we proposed a model-based mJSW measurement technique that could alleviate the problems related to tibial plateau alignment and patient positioning. The proposed technique resembles the one which is presented for TKAs in previous chapters. The main difference is that patient-specific bone geometries (the tibia and the femur) have to be reconstructed whereas the geometry of TKAs components is generally easy to obtain. To reconstruct patient-specific bone geometries without resorting to CT scans (due to i.e. radiation exposure and costs), statistical shape models were used. These models generate shapes by matching the edges of the femoral and tibial silhouettes found in planar radiographs, constrained by a likelihood condition of the expected shape learned from a training set of example shapes.

The validation in Chapter 7 showed that the smallest detectable difference in thickness was higher for the model-based measurement than for the conventional measurement (0.82 mm vs 0.70 mm). This means that the model-based measurement does not improve (early) wear detection. The predominant cause of this result was the modelling error resulting from the reconstruction of the bone structures with SSMs. Despite this finding, the results are encouraging for further research, which should focus on improvements in model and matching accuracy. For instance, advanced matching algorithms using multiscale information or edge orientation could be used to improve edge detecting and selection, leading to a higher precision [36, 117]. Also, the shape generation could be extended with non-linear shape deformation modules, which will reduce matching error especially when the SSM is too constrained to match unseen shapes [118]. In our opinion such improvements are feasible, thus encouraging for future work on the application of SSMs for native knees. The quantitative shape-analysis capabilities of SSMs can be highly valuable for both mJSW measurements *and* shape related osteoarthritis-research. Our model-based technique would then provide an economic alternative to MRI-based assessments [107, 119].

8-2 Future Work / Recommendations

This work presents convincing evidence that the knee mJSW measurement accuracy and precision is improved using model-based measurement techniques in RSA images as well as in standard AP radiographs. The next steps towards clinical application are to improve the measurement software and to conduct further research on the influence of knee flexion and implant design on the reliability of insert thickness measurements.

Measurement software

The current measurement software was a prototype adequate for experimental purposes. A single measurement takes several minutes for an experienced user and requires a cascade of different applications. For clinical practice, an integral application is required in which the measurement can be conducted within approximately 30 seconds and in a user-friendly manner. Especially relevant steps are the automation of the contour detection and visualization of the mJSW measurement. Integration in the existing model-based RSA software seems a good candidate since most of the analyses required are already at hand.

The model-based measurement software requires precise scanned 3D models of implant components to obtain reliable mJSW measurements. These models are not always available thus increasing the cost and complexity of this measurement compared to the conventional approach. Yet, it is expected that this disadvantage will diminish as the use of 3D models becomes more common in medicine, increasing cost efficiency. Optionally, the number of required scans can be reduced by using a single model per component size and type. Patient-specific component differences do exist (e.g. due to manufacturing tolerance and polishing of the components), but the influence is marginal.

The influence of flexion and implant design on measurement reliability

The findings from the retrieval study (Chapter 5) suggest that the loss of tibiofemoral contact or differences in tibiofemoral contact location influence the reliability of the measurement of the insert thickness based on the mJSW. Since reliability and accuracy of measurement are prerequisites for use in a clinical application, further research into this topic is necessary. This is closely related to the articulation pattern of a TKA and therefore knee flexion and implant design are important factors in this research.

For this research, using RSA instead of plain radiographs is recommended. The main reason is the more accurate reconstruction of the tibiofemoral contact location in

RSA due to the higher accuracy in the out-of-plain direction. This tibiofemoral contact location could be used as an indicator of measurement precision: In case implant designs are less congruent (i.e. more mobility at the articular surface), differences in flexion angles of the knee will cause large variations in contact location between femur and tibial insert. Due to this a lower precision of the mJSW measurement is expected to be present. Differences in contact location between successive measurements throughout follow-up could therefore indicate a limited measurement precision.

The findings from this research can be translated into conditions that should be met when conducting model-based mJSW measurements in RSA as well as in standard radiographs.

Model-based measurements as a diagnostic toolbox

In potential, model-based reconstructions and measurements allow for several diagnostics from a single image. Model-based techniques could thus provide a diagnostic toolbox for an integral, *in vivo* assessment of TKAs from radiographs. Examples of such an automated analysis are already found in the literature [120, 121]. For example, model-based RSA is already used to predict failure rates of new implant designs related to loosening. With addition of the model-based mJSW measurement, wear-related complications of total knee replacements could be predicted at the same time. Furthermore, model-based reconstructions could be used to model the bone geometry and herewith measure the alignment of the prosthetic components.

The main challenges of realising such an integrated toolbox are to reduce the processing workload using more automatic and faster procedures and to improve the model accuracy to obtain an acceptable measurement precision when deformable models are used. Given the rapid improvements in image quality, segmentation techniques and model accuracy as well as the fast developments in user friendliness, processing speed and reduction of costs of imaging software, we foresee that model-based wear measurements for native knees and TKAs will be common practice in the future.

References

[2] Hellio Le Graverand, M.-P., et al., *Radiographic grading and measurement of joint space width in osteoarthritis.* Rheumatic diseases clinics of North

America, 2009. **35**: p. 485-502.

[9] *Swedish Knee Registry, Anual report 2011*. 28-08-2012[cited 2012 August]; Available from: http://www.knee.nko.se/english/online/thePages/publication.php.

[10] Quinet, R.J. and E.G. Winters, *Total joint replacement of the hip and knee.* The Medical clinics of North America, 1992. **76**(5): p. 1235-51.

[11] Carr, A.J., et al., *Knee replacement.* The Lancet, 2012. **379**: p. 1331-1340.

[12] *Australian Orthopeadic Assocation National Joint Replacement Registry Anual report 2011*.

[13] *New Zealand National Joint Registry Annual Report 2009*.

[14] Robertsson, O., et al., *Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden.* Acta Orthopaedica Scandinavica, 2000. **71**: p. 262-7.

[15] Kurtz, S.M., et al., *Prevalence of primary and revision hip and knee arthroplasty in the United States from 1990 through 2002.* The Journal of Bone & Joint Surgery, 2005. **87**(7): p. 1487-1497.

[20] Sundfeldt, M., et al., *Aseptic loosening, not only a question of wear: a review of different theories.* Acta Orthopaedica, 2006. **77**: p. 177-97.

[23] Engh, C.A., et al., *Radiographically measured total knee wear is constant and predicts failure.* The Journal of Arthroplasty, 2013.

[25] Dennis, D.A. and R.D. Komistek, *Mobile-bearing total knee arthroplasty: design factors in minimizing wear.* Clinical Orthopaedics and Related Research, 2006. **452**: p. 70-7.

[27] Miller, T.T., *Imaging of knee arthroplasty.* European Journal of Radiology, 2005. **54**: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, *Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses.* The Journal of Arthroplasty, 2003. **18**: p. 860-866.

[29] Collier, M.B., et al., *Radiographic assessment of the thickness lost from polyethylene tibial inserts that had been sterilized differently.* The Journal of Bone & Joint Surgery. American Volume, 2008. **90**: p. 1543-52.

[31] Mintz, A.D., C.A. Pilkington, and D.W. Howie, *A comparison of plain and fluoroscopically guided radiographs in the asessment of arthroplasty of the*

knee. The Journal of Bone & Joint Surgery. American Volume, 1989. **71**: p. 1343-1347.

[33] Kaptein, B.L., et al., *Clinical validation of model-based RSA for a total knee prosthesis.* Clinical Orthopaedics and Related Research, 2007. **464**: p. 205-9.

[35] Vignon, E., et al., *Measurement of radiographic joint space width in the tibiofemoral compartment of the osteoarthritic knee: comparison of standing anteroposterior and Lyon schuss views.* Arthritis & Rheumatism, 2003. **48**: p. 378-84.

[36] Baka, N., et al., *2D-3D shape reconstruction of the distal femur from stereo X-Ray imaging using statistical shape models.* Medical Image Analysis, 2011. **15**(6): p. 840-850.

[41] Fukuoka, Y., A. Hoshino, and A. Ishida, *A simple radiographic measurement method for polyethylene wear in total knee arthroplasty.* IEEE Transactions of Rehabilitation Engineering, 1999. **7**: p. 228-233.

[42] Kellett, C.F., et al., *In vivo measurement of total knee replacement wear.* The Knee, 2004. **11**: p. 183-187.

[43] Short, A., et al., *A novel method for in vivo knee prosthesis wear measurement.* Journal of Biomechanics, 2005. **38**: p. 315-322.

[46] Nilsson, K.G. and J. Kärrholm, *RSA in the assessment of aseptic loosening.* The Journal of Bone & Joint Surgery. British Volume, 1996. **78**: p. 1-3.

[48] Gill, H.S., et al., *In vivo measurement of volumetric wear of a total knee replacement.* The Knee, 2006. **13**: p. 312-7.

[49] Simpson, D.J., et al., *Development of a model-based RSA systeem to measure polyethylene wear in unicompartmental arthroplasty.* Clinical Biomechanics, 2010. **?**: p.?

[65] Hoshino, A., A. Ishida, and Y. Fukuoka, *Accurate in vivo measurement of polyethylene wear in total knee arthroplasty.* The Journal of Arthroplasty, 2002. **17**: p. 490-496.

[79] Harman, M.K., et al., *Comparison of polyethylene tibial insert damage from in vivo function and in vitro wear simulation.* Journal of Orthopaedic Research, 2009. **27**(4): p. 540-548.

[82] Harman, M.K., S.A. Banks, and W.A. Hodge, *Polyethylene damage and knee kinematics after total knee arthroplasty.* Clinical Orthopaedics and Related

Research, 2001. **392**: p. 383-93.

[88] Guermazi, A., et al., *Imaging in osteoarthritis.* Rheumatic diseases clinics of North America, 2008. **34**: p. 645-87.

[100] Dupuis, D.E., et al., *Precision and accuracy of joint space width measurements of the medial compartment of the knee using standardized MTP semi-flexed radiographs.* Osteoarthritis and Cartilage, 2003. **11**(10): p. 716-724.

[102] Wirth, W., et al., *Direct comparison of fixed flexion, radiography and MRI in knee osteoarthritis: responsiveness data from the Osteoarthritis Initiative.* Osteoarthritis and Cartilage, 2013. **21**(1): p. 117-125.

[107] Felson, D.T. and T. Neogi, *Osteoarthritis: is it a disease of cartilage or of bone?* Arthritis & Rheumatism, 2004. **50**(2): p. 341-344.

[108] Skou, S.T., et al., *A Randomized, Controlled Trial of Total Knee Replacement.* New England Journal of Medicine, 2015. **373**(17): p. 1597-1606.

[109] Gascoyne, T.C., J.B. Morrison, and T.R. Turgeon, *The effect of patient positioning on the precision of model-based radiostereometric analysis.* Journal of Biomechanics, 2014. **47**(8): p. 1914-1917.

[110] Prins, A.H., et al., *Detecting condylar contact loss using single-plane fluoroscopy: A comparison with in vivo force data and in vitro bi-plane data.* Journal of Biomechanics, 2014. **47**(7): p. 1682-1688.

[111] Wasielewski, R.C., et al., *Wear patterns on retrieved polyethylene tibial inserts and their relationship to technical considerations during total knee arthroplasty.* Clinical Orthopaedics and Related Research, 1994. **299**: p. 31-43.

[112] Garling, E.H., *Objective clinical performance outcome of total knee prostheses. A study of mobile bearing knees using fluoroscopy, electromyography and roentgenstereophotogrammetry*, in *Department of Orthopaedics*2008, Leiden University Medical Center: Leiden.

[113] Teeter, M.G., et al., *Quantification of in vivo implant wear in total knee replacement from dynamic single plane radiography.* Physics in medicine and biology, 2013. **58**: p. 2751-67.

[114] Neumann, G., et al., *Location specific radiographic joint space width for osteoarthritis progression.* Osteoarthritis and Cartilage, 2009. **17**: p. 761-5.

[115] Kinds, M.B., et al., *Influence of variation in semiflexed knee positioning during image acquisition on separate quantitative radiographic parameters*

of osteoarthritis, measured by Knee Images Digital Analysis. Osteoarthritis and Cartilage. **20**(9): p. 997-1003.

[116] Hunter, D., et al., *OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis.* Osteoarthritis and Cartilage, 2015. **23**(5): p. 698-715.

[117] Prins, A., et al., *Integrated contour detection and pose estimation for fluoroscopic analysis of knee implants.* Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 2011. **225**(8): p. 753-761.

[118] Zheng, G., et al., *A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images.* Medical Image Analysis, 2009. **13**: p. 883-899.

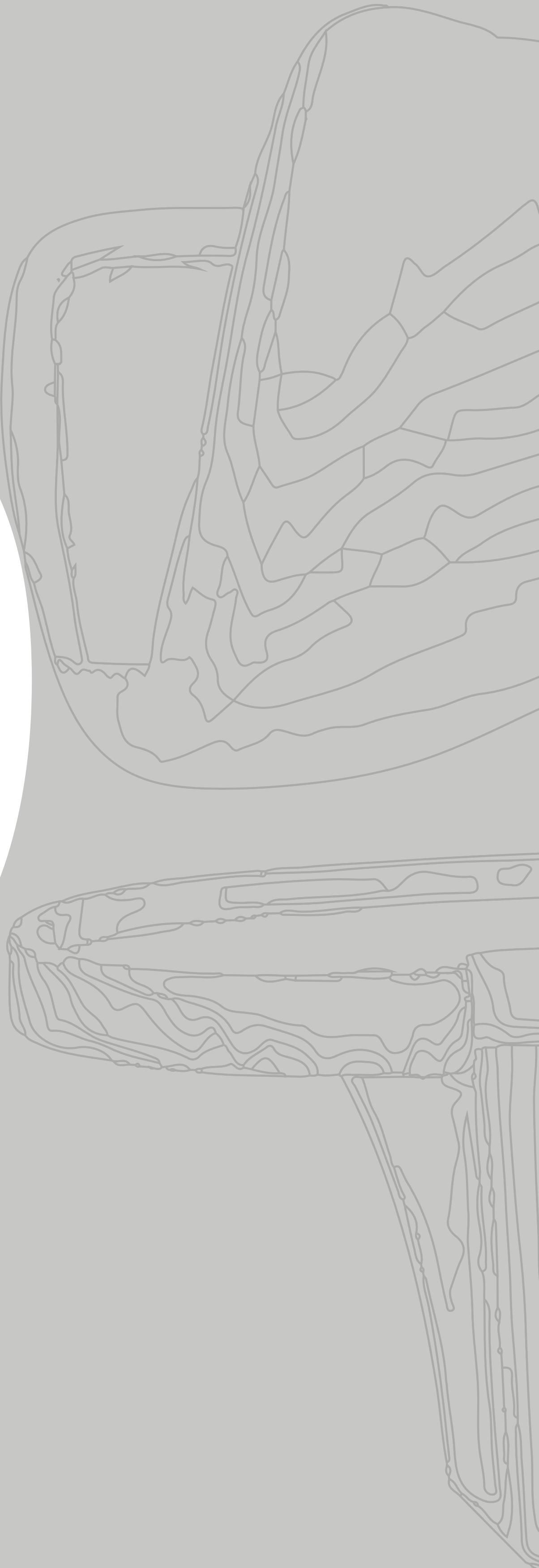
[119] Neogi, T., et al., *Magnetic Resonance Imaging–Based Three-Dimensional Bone Shape of the Knee Predicts Onset of Knee Osteoarthritis: Data From the Osteoarthritis Initiative.* Arthritis & Rheumatism, 2013. **65**(8): p. 2048-2058.

[120] Syu, C.-B., et al., *Automatic model-based roentgen stereophotogrammetric analysis (RSA) of total knee prostheses.* Journal of Biomechanics, 2012. **45**: p. 164-71.

[121] Kim, Y., et al., *Novel methods for 3D postoperative analysis of total knee arthroplasty using 2D-3D image registration.* Clinical Biomechanics, 2011. **26**: p. 384-91.

9

Summary



Summary

The primary aim of this work was to develop novel model-based mJSW measurement methods using a 3D reconstruction and validate the accuracy and precision of these methods. The model-based measurement results were compared to conventional mJSW measurement results. This thesis contributed to the development, validation and clinical application of model-based mJSW measurements for the natural knee and for total knee prostheses (TKAs).

The first chapters of this thesis (Chapter 2 and 3) focus on the in vitro validation of the model-based mJSW measurement for TKAs for model-based RSA and standard radiographs respectively. These studies showed that the model-based mJSW measurement is robust to variations in phantom positioning and prosthesis design. The best accuracy and precision was found for RSA (0.1 mm and 0.2 mm respectively). For standard radiographs, the accuracy and precision were superior to the conventional measurement methods.

The purpose of Chapter 4 was to investigate whether the measurement outcome is different in weight-bearing (WB) and non-weight-bearing (NWB) images due to knee laxity. This was investigated with 23 TKAs from an ongoing RSA study. The mJSW measured for the condyles was significantly larger in NWB images (difference of 0.28 mm medially and 0.20 mm laterally). In conclusion, mJSW are influences by

knee laxity for NWB images.

In Chapter 5 the validation of the model-based mJSW measurement for plain radiographs is continued with an in vivo study. The actual thickness of 15 retrieved inserts was compared to the mJSW measured in pre-operative radiographs. This study showed that the model-based measurement had a higher accuracy and a similar precision compared to the conventional measurement. It seems that the measurement outcome is influenced by differences in femoral contact location or loss of femoral contact.

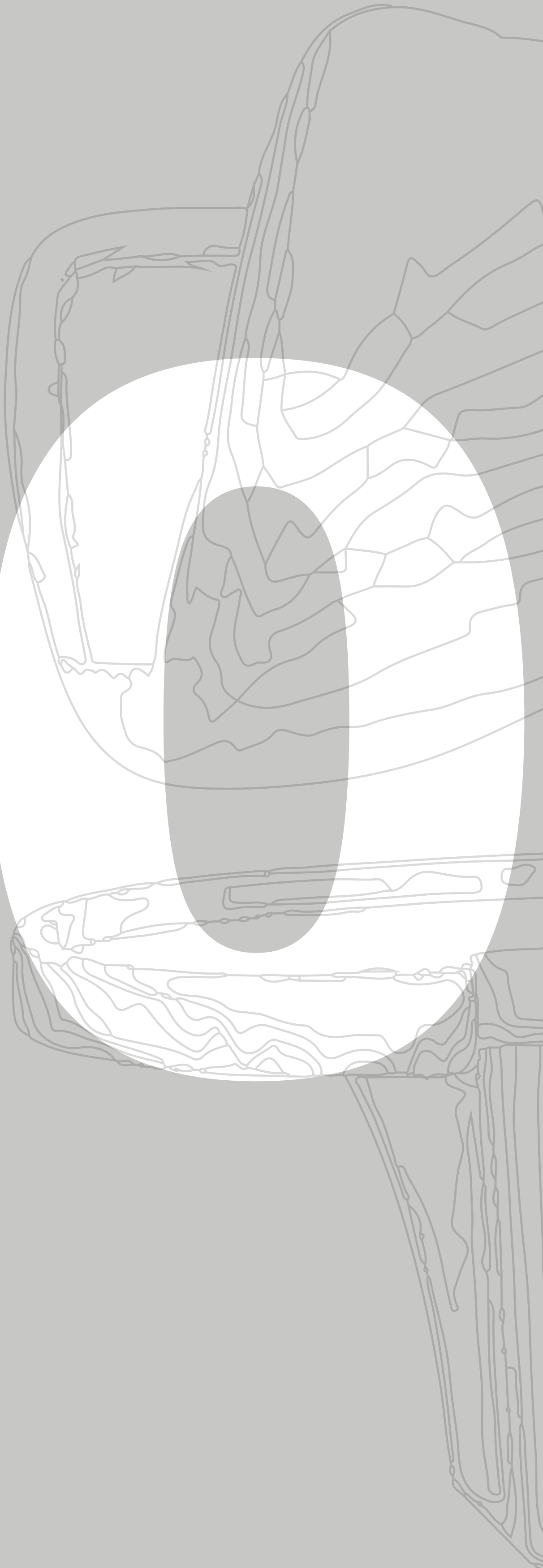
Model-based techniques can also be used to measure the TKA wear volume. In Chapter 6, simulations were conducted to assess the robustness of this measurement technique. The current error in 3D pose estimations in RSA imposes a considerable impact on the precision of volumetric wear measurements. The measurement was validated with inserts of which the wear volume was known. Results showed that at most 56% of the true wear volume was detected. The use of the measurement with the current technology is not recommended.

Chapter 7 shifts focus to the validation of mJSW measurements for the natural knee. In this case, statistical Shape Models (SSMs) are used to reconstruct the patient specific tibia and femur based on stereo images. It is shown that the SSM-based mJSW measurement method has a higher robustness but lower detectible mJSW difference than the conventional 2D method. Further research is required into improvements such as the use of a larger training set or smarter correspondence algorithms based on edge orientation or feature detection.

In conclusion, this work presents convincing evidence that the mJSW measurement accuracy and precision is improved using model-based measurement techniques in RSA images as well as in standard AP radiographs. The next steps towards clinical application are to improve the measurement software and to conduct further research on the influence of knee flexion and implant design on the reliability of mJSW as surrogate for the insert thickness.

10

Samenvatting



Samenvatting

Het primaire doel van dit werk was om nieuwe model-gebaseerde mJSW meetmethodes op basis van 3D-reconstructie te ontwikkelen en de nauwkeurigheid en precisie van deze methoden te valideren. De resultaten werden vergeleken met resultaten van conventionele mJSW methodes. Dit proefschrift draagt bij aan de ontwikkeling, validatie en klinische toepassing van modelgebaseerde mJSW metingen voor de natuurlijke knie en voor de totale knieprothesen (TKAs).

De eerste hoofdstukken van dit proefschrift (hoofdstuk 2 en 3) focussen op de in vitro validatie van de model-gebaseerde mJSW meting voor TKAs voor respectievelijk model-gebaseerde RSA en standaard röntgenfoto's. Deze studies toonden aan dat de modelgebaseerde mJSW meting robuust is voor variaties in fantoom positionering en prothese ontwerp. De beste nauwkeurigheid en precisie werd gevonden voor Röntgen Stereofotogrammetrische Analyse (RSA) (0,1 mm en 0,2 mm respectievelijk). Voor standaard röntgenfoto zijn de nauwkeurigheid en precisie superieur ten op zichte van de conventionel meetmethode.

Het doel van hoofdstuk 4 was om te onderzoeken of de meetuitkomst verschilt tussen belaste en niet-belaste beelden door laxiteit. Dit werd onderzocht met 23 TKAs uit een lopend RSA studie. De mJSW gemeten voor de condyles was significant groter in niet-belaste beelden (verschil van 0,28 mm mediaal en 0.20 mm lateraal). De conclusie uit dit hoofdstuk is dat de mJSW meting inderdaad wordt beinvloed door door laxiteit.

In hoofdstuk 5 wordt de validatie van de model-gebaseerde mJSW meting voor gewone röntgenfoto's voortgezet in een in vivo studie. De feitelijke dikte van 15 opgehaalde inserts werd vergeleken met de mJSW gemeten in pre-operatieve röntgenfoto's. De studie toonde aan dat het model-gebaseerde meting een hogere nauwkeurigheid en eenzelfde precisie heeft in vergelijking met de conventionele meting. Mogelijk wordt de meetuitkomst beïnvloed door verschillen in het femorale contactpunt of verlies van femoraal contact.

Model-gebaseerde technieken kunnen ook worden gebruikt om het volume van TKA slijtage te meten. In hoofdstuk 6 werden simulaties uitgevoerd om de robuustheid van deze meettechniek te beoordelen. De fout in het schatten van model posities en orientaties in RSA heeft een aanzienlijke invloed op de nauwkeurigheid van volumetrische slijtage metingen. De meting werd tevens gevalideerd met inserts waarvan het slijtagevolume bekend was. De resultaten toonden aan dat ten hoogste 56% van het werkelijke slijtagevolume gedetecteerd. Het gebruik van de meting met de huidige technologie wordt daarom vooralsnog afgeraden.

Hoofdstuk 7 gaat in op de validatie van mJSW metingen voor de natuurlijke knie. In dit geval worden Statistical Shapemodels (SSM) om de patiënt-specifieke tibia en femur te reconstrueren op basis van stereobeelden. Aangetoond wordt dat de SSM-gebaseerde mJSW meetmethode heeft een hogere robuustheid maar lager waarneembaar mJSW verschil dan de conventionele 2D-methode. Verder onderzoek is nodig naar mogelijke verbeteringen, zoals het gebruik van een grotere training set of slimmere algoritmen die model-beeld correspondentie zoeken gebaseerd op de contour orientatie of feature detectie.

Samenvattend toont dit werk overtuigend bewijs dat de nauwkeurigheid en precisie van mJSW metingen verbeteren door het gebruik van model-gebaseerde meettechnieken in zowel RSA beelden als in standaard AP röntgenfoto. De volgende stappen naar klinische toepassing zijn het verbeteren van de meetsoftware en het verder onderzoeken van de invloed van flexie en TKA vormverschillen op de betrouwbaarheid van de mJSW als maat voor de insert dikte.

List of bibliography

[1] Shapiro, J.R., et al., Osteogenesis imperfecta: A translational approach to brittle bone disease2013: Academic Press.

[2] Hellio Le Graverand, M.-P., et al., Radiographic grading and measurement of joint space width in osteoarthritis. Rheumatic diseases clinics of North America, 2009. 35: p. 485-502.

[3] Poos, M.J.J.C. and A.M. Gommer. Hoe vaak komt arthrose voor en hoeveel mensen sterven er aan? 2014 23 juni 2014; Available from: http://www.nationaalkompas.nl/gezondheid-en-ziekte/ziekten-en-aandoeningen/bewegingsstelsel-en-bindweefsel/artrose.

[4] RIVM, Kosten van Ziektenstudie, 2013, RIVM.

[5] Leskinen, J., et al., The incidence of knee arthrolasty for primary osteoarthritis grows rapidly among baby-boomers - a population-based study. Arthritis & Rheumatism, 2012. 64: p. 423-428.

[6] Kurtz, S.M., et al., Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. Clinical Orthopaedics and Related Research, 2009. 467: p. 2606-12.

[7] Murphy, L., et al., Lifetime risk of symptomatic knee osteoarthritis. Arthritis Care & Research, 2008. 59(9): p. 1207-1213.

[8] Wyatt, S.B., K.P. Winters, and P.M. Dubbert, Overweight and obesity: prevalence, consequences, and causes of a growing public health problem.

The American Journal of the Medical Sciences, 2006. 331(4): p. 166-74.
[9] Swedish Knee Registry, Anual report 2011. 28-08-2012 [cited 2012 August]; Available from: http://www.knee.nko.se/english/online/thePages/publication.php.
[10] Quinet, R.J. and E.G. Winters, Total joint replacement of the hip and knee. The Medical clinics of North America, 1992. 76(5): p. 1235-51.
[11] Carr, A.J., et al., Knee replacement. The Lancet, 2012. 379: p. 1331-1340.
[12] Australian Orthopeadic Assocation National Joint Replacement Registry Anual report 2011.
[13] New Zealand National Joint Registry Annual Report 2009.
[14] Robertsson, O., et al., Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. Acta Orthopaedica Scandinavica, 2000. 71: p. 262-7.
[15] Kurtz, S.M., et al., Prevalence of primary and revision hip and knee arthroplasty in the United States from 1990 through 2002. The Journal of Bone & Joint Surgery, 2005. 87(7): p. 1487-1497.
[16] Naudie, D.D.R., et al., Wear and osteolysis around total knee arthroplasty. The Journal of the American Academy of Orthopaedic Surgeons, 2007. 15: p. 53-64.
[17] Schmalzried, T.P. and J.J. Callaghan, Current Concepts Review-Wear in Total Hip and Knee Replacements. The Journal of Bone & Joint Surgery, 1999. 81(1): p. 115-136.
[18] Lewis, G., Polyethylene wear in total hip and knee arthroplasties. Journal of Biomedical Materials Research, 1997. 38: p. 55-75.
[19] Ingham, E. and J. Fisher, The role of macrophages in osteolysis of total joint replacement. Biomaterials, 2005. 26(11): p. 1271-86.
[20] Sundfeldt, M., et al., Aseptic loosening, not only a question of wear: a review of different theories. Acta Orthopaedica, 2006. 77: p. 177-97.
[21] Mulhall, K.J., et al., Current etiologies and modes of failure in total knee arthroplasty revision. Clinical Orthopaedics and Related Research, 2006. 446: p. 45-50.
[22] Sharkey, P.F., et al., Why are total knee arthroplasties failing today? Clinical Orthopaedics and Related Research, 2002. 404: p. 7-13.
[23] Engh, C.A., et al., Radiographically measured total knee wear is constant and

predicts failure. The Journal of Arthroplasty, 2013.

[24] Pijls, B.G., H.M. Van der Linden-Van der Zwaag, and R.G. Nelissen, Polyethylene thickness is a risk factor for wear necessitating insert exchange. International Orthopaedics, 2012. 36(6): p. 1175-80.

[25] Dennis, D.A. and R.D. Komistek, Mobile-bearing total knee arthroplasty: design factors in minimizing wear. Clinical Orthopaedics and Related Research, 2006. 452: p. 70-7.

[26] Lavernia, C.J., et al., Activity level and wear in total knee arthroplasty: a study of autopsy retrieved specimens. The Journal of Arthroplasty, 2001. 16: p. 446-53.

[27] Miller, T.T., Imaging of knee arthroplasty. European Journal of Radiology, 2005. 54: p. 164-77.

[28] Collier, M.B., B.A. Jewett, and C.A. Engh, Clinical assessment of tibial polyethylene thickness comparison of radiographic measurements with as-implanted and as-retrieved thicknesses. The Journal of Arthroplasty, 2003. 18: p. 860-866.

[29] Collier, M.B., et al., Radiographic assessment of the thickness lost from polyethylene tibial inserts that had been sterilized differently. The Journal of Bone & Joint Surgery. American Volume, 2008. 90: p. 1543-52.

[30] Sanzén, L., et al., Radiographic wear assessment in a total knee prosthesis. The Journal of Arthroplasty, 1996. 11: p. 738-42.

[31] Mintz, A.D., C.A. Pilkington, and D.W. Howie, A comparison of plain and fluoroscopically guided radiographs in the asessment of arthroplasty of the knee. The Journal of Bone & Joint Surgery. American Volume, 1989. 71: p. 1343-1347.

[32] Kaptein, B.L., et al., A new model-based RSA method validated using CAD models and models from reversed engineering. Journal of Biomechanics, 2003. 36: p. 873-882.

[33] Kaptein, B.L., et al., Clinical validation of model-based RSA for a total knee prosthesis. Clinical Orthopaedics and Related Research, 2007. 464: p. 205-9.

[34] Ryd, L., et al., Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses. The Journal of Bone & Joint Surgery. British Volume, 1995. 77: p. 377-83.

[35] Vignon, E., et al., Measurement of radiographic joint space width in the

tibiofemoral compartment of the osteoarthritic knee: comparison of standing anteroposterior and Lyon schuss views. Arthritis & Rheumatism, 2003. 48: p. 378-84.

[36] Baka, N., et al., 2D-3D shape reconstruction of the distal femur from stereo X-Ray imaging using statistical shape models. Medical Image Analysis, 2011. 15(6): p. 840-850.

[37] Baka, N., et al., Correspondence free 3D statistical shape model fitting to sparse x-ray projections. SPIE, 2010: p. 76230D-76230D-9.

[38] Baldwin, M.a., et al., Development of subject-specific and statistical shape models of the knee using an efficient segmentation and mesh-morphing approach. Computer methods and programs in Biomedicine, 2010. 97: p. 232-40.

[39] Burckhardt, K., et al., Measuring orthopedic implant wear on standard radiographs with a precision in the 10 micrometer range. Medical Image Analysis, 2006. 10: p. 520-529.

[40] Duryea, J., et al., Fully automated software to monitor wear in prosthetic knees using fluoroscopic images. European Radiology, 2001. 11: p. 2184-7.

[41] Fukuoka, Y., A. Hoshino, and A. Ishida, A simple radiographic measurement method for polyethylene wear in total knee arthroplasty. IEEE Transactions of Rehabilitation Engineering, 1999. 7: p. 228-233.

[42] Kellett, C.F., et al., In vivo measurement of total knee replacement wear. The Knee, 2004. 11: p. 183-187.

[43] Short, A., et al., A novel method for in vivo knee prosthesis wear measurement. Journal of Biomechanics, 2005. 38: p. 315-322.

[44] Garling, E.H., E.R. Valstar, and R.G.H.H. Nelissen, Comparison of micromotion in mobile bearing and posterior stabilized total knee prostheses: a randomized RSA study of 40 knees followed for 2 years. Acta Orthopaedica, 2005. 76: p. 353-61.

[45] Nelissen, R.G.H.H., Fundamental aspects of the clinication evalution of total knee prostheses, 1995.

[46] Nilsson, K.G. and J. Kärrholm, RSA in the assessment of aseptic loosening. The Journal of Bone & Joint Surgery. British Volume, 1996. 78: p. 1-3.

[47] Soballe, K., et al., Migration of hydroxyapatite coated femoral prostheses. A Roentgen Stereophotogrammetric study. The Journal of Bone & Joint Surgery.

British Volume, 1993. 75: p. 681.
[48] Gill, H.S., et al., In vivo measurement of volumetric wear of a total knee replacement. The Knee, 2006. 13: p. 312-7.
[49] Simpson, D.J., et al., Development of a model-based RSA systeem to measure polyethylene wear in unicompartmental arthroplasty. Clinical Biomechanics, 2010. ?: p.?
[50] Kaptein, B.L., et al., Evaluation of three pose estimation algorithms for model-based Roentgen stereophotogrammetric analysis, in The Journal of Engineering in Medicine, part H2004. p. 231-8.
[51] Bartel, D.L., V.L. Bicknell, and T.M. Wright, The effect of conformity, thickness, and material on stresses in ultra-high molecular weight components for total joint replacement. The Journal of Bone & Joint Surgery. American Volume, 1986. 68: p. 1041-1051.
[52] Deng, M., et al., Study of creep behavior of ultra-high-molecular-weight polyethylene systems. Journal of Biomedical Materials Research, 1998. 40: p. 214-223.
[53] Lee, K.-Y. and D. Pienkowski, Compressive creep characteristics of extruded ultrahigh-molecular-weight polyethylene. Journal of Biomedical Materials Research, 1998. 39: p. 261-265.
[54] Robertsson, O., et al., Past Incidence and future demands for knee arthroplasty in Sweden. Acta Orthopaedica, 2000. 71: p. 376-380.
[55] Bert, J.M., et al., The incidence of modular tibial polyethylene insert exchange in total knee arthroplasty when polyethylene failure occurs. The Journal of Arthroplasty, 1998. 13: p. 609-14.
[56] Babis, G.C., R.T. Trousdale, and B.F. Morrey, The effectiveness of isolated tibial insert exchange in revision total knee arthroplasty. The Journal of Bone & Joint Surgery. American Volume, 2002. 84-A: p. 64-8.
[57] Engh, G.a., L.M. Koralewicz, and T.R. Pereles, Clinical results of modular polyethylene insert exchange with retention of total knee arthroplasty components. The Journal of Bone & Joint Surgery. American Volume, 2000. 82: p. 516-23.
[58] IJsseldijk, E.A.v., et al., The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA. Journal of Biomechanics, 2011. 44(15): p. 2724-2727.

[59] Wunsch, P. and G. Hirzinger, Registration of CAD-models to images by iterative inverse perspective matching. Proceedings of 13th International Conference on Pattern Recognition1996: IEEE. 78-83 vol.1.

[60] Canny, J., A Computational Approach To Edge Detection. Pattern Analysis and Machine Intelligence, IEEE transactions on, 1986. PAMI-8(6): p. 679-698.

[61] PITMAN, E.J.G., A note on normal correlation. Biometrika, 1939. 31: p. 9-12.

[62] Bland, J.M. and D.G. Altman, STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. The Lancet, 1986. 327(8476): p. 307-310.

[63] Kaptein, B.L., et al., Evaluation of three pose estimation algorithms for model-based Roentgen stereophotogrammetric analysis. J. Engineering in Medicine, part H, 2004: p. 231-8.

[64] Hide, I.G., et al., A radiological technique for the assessment of wear in prosthetic knee replacements. Skeletal Radiology, 2000. 29: p. 583-586.

[65] Hoshino, A., A. Ishida, and Y. Fukuoka, Accurate in vivo measurement of polyethylene wear in total knee arthroplasty. The Journal of Arthroplasty, 2002. 17: p. 490-496.

[66] Chapman-Sheath, P.J., et al., In vitro assessment of proximal polyethylene contact surface areas and stresses in mobile bearing knees. Medical Engineering & Physics, 2003. 25: p. 437-443.

[67] McEwen, H.M.J., et al., The influence of design, materials and kinematics on the in vitro wear of total knee replacements. Journal of Biomechanics, 2005. 38: p. 357-65.

[68] Fregly, B.J., et al., Computational wear prediction of a total knee replacement from in vivo kinematics. Journal of Biomechanics, 2005. 38: p. 305-14.

[69] Collier, M.B., et al., Factors associated with the loss of thickness of polyethylene tibial bearings after knee arthroplasty. The Journal of Bone & Joint Surgery. American Volume, 2007. 89: p. 1306-14.

[70] Nouta, K.A., et al., All-polyethylene tibial components are equal to metal-backed components: systematic review and meta-regression. Clinical Orthopaedics and Related Research, 2012. 470(12): p. 3549-59.

[71] Blagojevic, M., et al., Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthritis and Cartilage, 2010. 18(1): p. 24-33.

[72] van IJsseldijk, E.A., et al., The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA. Journal of Biomechanics, 2011. 44: p. 2724-7.

[73] Knight, J.L., R.D. Atwater, and J. Guo, Early failure of the porous coated anatomic cemented unicompartmental knee arthroplasty. Aids to diagnosis and revision. The Journal of Arthroplasty, 1997. 12: p. 11-20.

[74] Buckland-Wright, C., Which radiographic techniques should we use for research and clinical practice? Best Practice & Research in Clinical Rheumatology, 2006. 20: p. 39-55.

[75] Bragdon, C.R., et al., Standing versus supine radiographs in RSA evaluation of femoral head penetration. Clinical Orthopaedics and Related Research, 2006. 448: p. 46-51.

[76] van IJsseldijk, E.A., et al., A model-based approach to measure the minimum joint space width of total knee replacements in standard radiographs. Journal of Biomechanics, 2012. 45: p. 2171-5.

[77] Delport, H., et al., A kinematic comparison of fixed-and mobile-bearing knee replacements. The Journal of Bone & Joint Surgery. British Volume, 2006. 88(8): p. 1016-1021.

[78] Harman, M., et al., A pictographic atlas for classifying damage modes on polyethylene bearings. Journal of Materials Science: Materials in Medicine, 2011. 22(5): p. 1137-1146.

[79] Harman, M.K., et al., Comparison of polyethylene tibial insert damage from in vivo function and in vitro wear simulation. Journal of Orthopaedic Research, 2009. 27(4): p. 540-548.

[80] Harman, M.K., S.A. Banks, and W.A. Hodge, Backside damage corresponding to articular damage in retrieved tibial polyethylene inserts. Clinical Orthopaedics and Related Research, 2007. 458: p. 137-144.

[81] Martin Bland, J. and D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet, 1986. 327: p. 307-310.

[82] Harman, M.K., S.A. Banks, and W.A. Hodge, Polyethylene damage and knee kinematics after total knee arthroplasty. Clinical Orthopaedics and Related Research, 2001. 392: p. 383-93.

[83] Prins, A., et al., Detecting femur–insert collisions to improve precision of

fluoroscopic knee arthroplasty analysis. Journal of Biomechanics, 2010. 43(4): p. 694-700.

[84] Psychoyios, V., et al., Wear of congruent meniscal bearings in unicompartmental knee arthroplasty. The Journal of Bone & Joint Surgery. British Volume, 1998. 80: p. 976-982.

[85] Edwards, S.a., et al., Analysis of polyethylene thickness of tibial components in total knee replacement. The Journal of Bone & Joint Surgery. American Volume, 2002. 84-A: p. 369-71.

[86] Schaap, L.A., et al., European Project on OSteoArthritis (EPOSA): methodological challenges in harmonization of existing data from five European population-based cohorts on aging. BMC musculoskeletal disorders, 2011. 12(1): p. 272.

[87] Guermazi, A., et al., Imaging in osteoarthritis. Rheumatic diseases clinics of North America, 2008. 34: p. 645-87.

[88] Mazzuca, S.A., et al., Effect of alignment of the medial tibial plateau and X-ray beam on apparent progression of osteoarthritis in the standing anteroposterior knee radiograph. Arthritis & Rheumatism, 2001. 44: p. 1786-1794.

[89] Conrozier, T., et al., Influence of the quality of tibial plateau alignment on the reproducibility of computer joint space measurement from Lyon schuss radiographic views of the knee in patients with knee osteoarthritis. Osteoarthritis and Cartilage, 2004. 12: p. 765-70.

[90] Mazzuca, S.A., et al., Performance of a non-fluoroscopically assisted substitute for the Lyon schuss knee radiograph: quality and reproducibility of positioning and sensitivity to joint space narrowing in osteoarthritic knees. Osteoarthritis and Cartilage, 2008. 16: p. 1555-9.

[91] Buckland-Wright, J., et al., Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. The Journal of Rheumatology, 1999. 26(12): p. 2664-2674.

[92] Eckstein, W., E.B. Kley, and A. Tunnermann, Comparison of different simulation methods for effective medium computer generated holograms. Optics express, 2013. 21(10): p. 12424-33.

[93] Eckstein, F., T. Mosher, and D. Hunter, Imaging of knee osteoarthritis: data beyond the beauty. Current Opinion in Rheumatology, 2007. 19: p. 435-43.

[94] Felson, D.T. and T. Neogi, Osteoarthritis: is it a disease of cartilage or of bone?

Arthritis & Rheumatism, 2004. 50(2): p. 341-344.
[95] Sarkalkan, N., H. Weinans, and A.A. Zadpoor, Statistical shape and appearance models of bones. Bone, 2014. 60: p. 129-40.
[96] Huétink, K., et al., Automatic radiographic quantification of hand osteoarthritis; accuracy and sensitivity to change in joint space width in a phantom and cadaver study. Skeletal Radiology, 2012. 41: p. 41-9.
[97] Lorensen, W.E. and H.E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. in ACM Siggraph Computer Graphics. 1987. ACM.
[98] Cootes, T.F., et al., Use of active shape models for locating structures in medical images. Image and vision computing, 1994. 12(6): p. 355-365.
[99] Baka, N., et al., Evaluation of automated statistical shape model based knee kinematics from biplane fluoroscopy. Journal of Biomechanics, 2014. 47(1): p. 122-129.
[100] Klein, S., et al., Elastix: a toolbox for intensity-based medical image registration. Medical Imaging, IEEE Transactions on, 2010. 29(1): p. 196-205.
[101] Bland, J.M. and D.G. Altman, Statistics notes: measurement error. Bmj, 1996. 313(7059): p. 744.
[102] Dupuis, D.E., et al., Precision and accuracy of joint space width measurements of the medial compartment of the knee using standardized MTP semi-flexed radiographs. Osteoarthritis and Cartilage, 2003. 11(10): p. 716-724.
[103] Skou, S.T., et al., A Randomized, Controlled Trial of Total Knee Replacement. New England Journal of Medicine, 2015. 373(17): p. 1597-1606.
[104] Gascoyne, T.C., J.B. Morrison, and T.R. Turgeon, The effect of patient positioning on the precision of model-based radiostereometric analysis. Journal of Biomechanics, 2014. 47(8): p. 1914-1917.
[105] Prins, A.H., et al., Detecting condylar contact loss using single-plane fluoroscopy: A comparison with in vivo force data and in vitro bi-plane data. Journal of Biomechanics, 2014. 47(7): p. 1682-1688.
[106] Wasielewski, R.C., et al., Wear patterns on retrieved polyethylene tibial inserts and their relationship to technical considerations during total knee arthroplasty. Clinical Orthopaedics and Related Research, 1994. 299: p. 31-43.
[107] Garling, E.H., Objective clinical performance outcome of total knee prostheses. A study of mobile bearing knees using fluoroscopy, electromyography and roentgenstereophotogrammetry, 2008 (Dr). Department of Orthopaedics,

Leiden University Medical Center.

[108] Teeter, M.G., et al., Quantification of in vivo implant wear in total knee replacement from dynamic single plane radiography. Physics in medicine and biology, 2013. 58: p. 2751-67.

[109] Neumann, G., et al., Location specific radiographic joint space width for osteoarthritis progression. Osteoarthritis and Cartilage, 2009. 17: p. 761-5.

[110] Kinds, M.B., et al., Influence of variation in semiflexed knee positioning during image acquisition on separate quantitative radiographic parameters of osteoarthritis, measured by Knee Images Digital Analysis. Osteoarthritis and Cartilage. 20(9): p. 997-1003.

[111] Hunter, D., et al., OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis. Osteoarthritis and Cartilage, 2015. 23(5): p. 698-715.

[112] Wirth, W., et al., Direct comparison of fixed flexion, radiography and MRI in knee osteoarthritis: responsiveness data from the Osteoarthritis Initiative. Osteoarthritis and Cartilage, 2013. 21(1): p. 117-125.

[113] Prins, A., et al., Integrated contour detection and pose estimation for fluoroscopic analysis of knee implants. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 2011. 225(8): p. 753-761.

[114] Zheng, G., et al., A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. Medical Image Analysis, 2009. 13: p. 883-899.

[115] Neogi, T., et al., Magnetic Resonance Imaging–Based Three-Dimensional Bone Shape of the Knee Predicts Onset of Knee Osteoarthritis: Data From the Osteoarthritis Initiative. Arthritis & Rheumatism, 2013. 65(8): p. 2048-2058.

[116] Syu, C.-B., et al., Automatic model-based roentgen stereophotogrammetric analysis (RSA) of total knee prostheses. Journal of Biomechanics, 2012. 45: p. 164-71.

[117] Kim, Y., et al., Novel methods for 3D postoperative analysis of total knee arthroplasty using 2D-3D image registration. Clinical Biomechanics, 2011. 26: p. 384-91.

Publications

Journal Papers

- **van IJsseldijk EA**, Valstar ER, Stoel BC, Nelissen RGHH, Reiber JHC, Kaptein BL. 2011. The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA. Journal of Biomechanics 44:2724-2727.
- **van IJsseldijk EA**, Valstar ER, Stoel BC, Nelissen RGHH, Kaptein BL. 2012. A model-based approach to measure the minimum joint space width of total knee replacements in standard radiographs. Journal of Biomechanics 45:2171-2175.
- **van IJsseldijk EA**, Lebel B, Stoel BC, Valstar ER, Gouzy S, Vielpeau C, Kaptein BL. 2013. Validation of the in vivo volumetric wear measurement for total knee prostheses in model-based RSA. Journal of Biomechanics 46:1387-1391.
- **van IJsseldijk EA**, Valstar ER, Stoel BC, de Ridder R, Nelissen RGHH, Kaptein BL. 2014. Measuring polyethylene wear in total knee arthroplasty by RSA: differences between weight-bearing and non-weight-bearing positioning. Journal of Orthopaedic Research 32:613-617.
- **van IJsseldijk EA**, Harman MK, Luetzner J, Valstar ER, Stoel BC, Nelissen RGHH, Kaptein BL. 2014. Validation of a model-based measurement of the minimum insert thickness of knee prostheses. Bone and Joint Research 3(10):289-296
- **van IJsseldijk EA**, Valstar ER, Stoel BC, Nelissen RGHH, Baka N, van 't Klooster R, Kaptein BL. 2016. Three dimensional measurement of minimum joint space

width in the knee from stereo radiographs using statistical shape models, Bone and Joint Research, In press

Conference Abstracts

- **van IJsseldijk EA**, Valstar ER, Stoel BC, Nelissen RGHH, Reiber JHC, Kaptein BL, The Robustness and Accuracy of in vivo Linear Wear Measurements for Knee Prostheses based on Model-Based RSA, *RSA Meeting 2011*, Leiden, The Netherlands
- **van IJsseldijk EA**, Valstar ER, Stoel BC, Nelissen RGHH, Reiber JHC, Kaptein BL, A Model Based in vivo Linear Wear Measurement for Knee Prostheses in standard radiographs, *European Orthopaedic Research Society meeting 2011*, Vienna, Austria
- **van IJsseldijk EA**, Gouzy S, Lebel B, Kaptein BL, Valstar ER, Stoel BC, RGHH Nelissen, Vielpeau C, Sensitivity of total knee volumetric wear measurement using RSA, *European Orthopaedic Research Society meeting 2013*, Amsterdam, The Netherlands
- **van IJsseldijk EA**, Valstar ER, Stoel BC, Nelissen RGHH, Kaptein BL. Measuring polyethylene wear of total knee replacements, do we need bearing images?, *RSA Meeting 2013,* Lund, Sweden
- **van IJsseldijk EA**, Gouzy S, Lebel B, Kaptein BL, Valstar ER, Stoel BC, RGHH Nelissen, Vielpeau C, Validation of the in vivo volumetric wear measurement for total knee prostheses in model-based RSA, *RSA Meeting 2013*, Lund, Sweden (Poster presentation)

Curriculum Vitae

Emiel van IJsseldijk was born in 1981 in Schiedam. He studied Applied Physics at the Technical University of Delft (TU Delft), where he obtained his Bachelor of Science in 2005. He continued his education at TU Delft, which was completed with a Master of Science degree in Biomedical Engineering with a specialization in Medical Imaging in 2009. In 2007, he did an internship at BrainLAB AG (München, Germany). During this internship, he worked on algorithm development, testing and validation on the subject of leg length and offset calculation for navigated total hip replacement surgery. He developed and validated Statistical Shape Models of the scapula at the Quantitative Imaging Group of the faculty of Applied Physics of TU Delft for his Master thesis.

From June 2009 onwards, he pursued his interest in medical technology as a PhD student on the topic of applied model-based techniques for measurements in the knee and knee prostheses with this thesis as a result. A part of the research was conducted in collaboration with research groups in different in the USA, France and Germany.

Since 2013 Emiel is employed as a stress test expert at Rabobank Nederland in Utrecht. Here, he combines his skills in software development, data analysis and modelling in the financial sector.

Dankwoord

Ik kijk met veel plezier en dankbaarheid terug op dit promotietraject. Een traject waarbij ik door de jaren heen van diverse kanten ondersteuning heb gekregen van de mensen om mij heen.

Op inhoudelijk gebied heb ik veel profijt gehad van alle ideeën en suggesties van mijn promotoren. Met name die ideeën en suggesties op het moment dat ik zelf vooral ruis kon produceren. Ook ben ik ontzettend dankbaar voor de mogelijkheden die ik heb gekregen om mijn onderzoek in het buitenland te mogen uitvoeren en de prettige samenwerking die ik hierbij heb ervaren. Ik ben er zeker van dat dit een belangrijke motivatie is geweest om dit werk tot een goed einde te brengen.

Een technisch promotietraject kan je ook doen met twee linkerhanden. Ik wil daarom graag iedereen bedanken voor hun inzet en handigheidjes bij de voorbereiding en uitvoering van alle experimenten in het LUMC. Mijn dank gaat ook uit naar Andre, Nora en Ronald voor het tot mijn beschikking stellen en afstellen van hun software op mijn onderzoek. Verder wil ik speciaal Anika nog bedanken voor de structuur en ondersteuning die ze heeft geboden bij alle reglementaire zaken.

Aan mijn mede-promovendi: bedankt voor de mooie herinneringen. Zowel aan de bende die we onderzoekskamer mochten noemen als aan onze ervaringen elders in

Europa. Ik ben vereerd dat ik een aantal van jullie inmiddels tot mijn vrienden mag rekenen.

Marinda, ik kan niet precies in woorden uitdrukken hoeveel je hebt betekend voor mijn promotieonderzoek. Bedankt voor al je inzet, je goede raad, je geduld naar mij toe, maar zeker ook voor je ongeduld als het om resultaten gaat. Ik vraag me weleens af of ik anders ooit bij dit hoofdstuk was gekomen.

Tot slot gaat mijn dank ook uit naar mijn huidige collega's voor de ruimte en de waardering die ik krijg. Beide hebben geholpen om dit werk tot een goed einde te brengen.