TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

RADIOSTEREOMETRIC ANALYSIS IN TOTAL HIP ARTHROPLASTY AND HIP FRACTURE PATIENTS

Sami Finnilä

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1583 | MEDICA – ODONTOLOGICA | TURKU 2021

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

RADIOSTEREOMETRIC ANALYSIS IN TOTAL HIP ARTHROPLASTY AND HIP FRACTURE PATIENTS

Sami Finnilä

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1583 | MEDICA – ODONTOLOGICA | TURKU 2021

University of Turku

Faculty of Medicine
Orthopaedics and Traumatology
Doctoral Programme in Clinical Research

Supervised by

Prof. Emeritus Hannu T Aro, MD, PhD
Department of Orthopaedic Surgery and Traumatology,
University of Turku and Turku University Hospital, Finland

Adjunct Prof. Niko Moritz, PhD
Turku Clinical Biomaterial Centre and Institute of Dentistry,
University of Turku, Finland

Reviewed by

Prof. Miika Nieminen, PhD
Research Unit of Medical Imaging, Physics and Technology,
University of Oulu, Finland

Adjunct Prof. Teemu Moilanen, MD, PhD
Department of Orthopaedic Surgery and Traumatology,
University of Tampere, Finland

Opponent

Prof. Juhana Leppilahti, MD, PhD
Department of Orthopaedics and Traumatology,
Oulu University Hospital, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-8619-4 (PRINT)
ISBN 978-951-29-8620-0 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama Oy, Turku, Finland, 2021

To my family: Tiina and Sonja

UNIVERSITY OF TURKU
Faculty of Medicine
Orthopaedics and Traumatology
SAMI FINNILÄ: Radiostereometric Analysis in Total Hip Arthroplasty and Hip Fracture Patients
Doctoral thesis, 178 pp.
Doctoral Programme in Clinical Research (DPCR)
October 2021

ABSTRACT

Complications related to primary total hip arthroplasty (THA) are relatively rare but still impose a significant burden on the recovery of individual patients and incur significant costs to the healthcare system. Research aimed at improving the results of THA is challenging as complications can take up to decades to manifest clinically. However, radiostereometric analysis (RSA) can, in some cases, be used to predict the long-term revision rates of THA with only a two-year follow-up. The purpose of this doctoral thesis was to examine the causes of RSA-measured micromotion and to further develop the methodology for the research of THA and hip fracture patients.

The first study examined whether preoperative systemic bone mineral density (BMD) had an effect on the early RSA-measured micromotion of a cementless acetabular cup in female patients with osteoarthritis. The second study considered the suitability of model-based RSA (MBRSA) for the analysis of a cementless femoral stem using both a phantom model and a clinical cohort. The third study validated differentially-loaded RSA (DLRSA) for the study of internally-fixated femoral neck fractures in a clinical cohort of 16 patients. The final study examined if RSA data analysis would benefit from the use of a multivariate three-dimensional analytical method.

Low systemic BMD was associated with increased proximal migration of the cementless acetabular cups. The MBRSA proved to have comparable accuracy and precision compared to conventional RSA thereby validating the method for future clinical studies using the examined femoral stem. The deployed DLRSA methodology could be used to detect inducible micromotion of femoral neck fractures. A multivariate linear mixed-effects model could provide a more robust and sensitive method for the analysis of three-dimensional RSA data.

KEYWORDS: Radiostereometric analysis, total hip arthroplasty, femoral neck fracture, osteoporosis, three-dimensional micromotion

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Ortopedia ja traumatologia
SAMI FINNILÄ: Radiostereometrinen analyysi lonkan kokotekonivelen ja lonkkamurtumien tutkimuksessa
Väitöskirja, 178 s.
Turun kliininen tohtoriohjelma (TKT)
October 2021

TIIVISTELMÄ

Lonkan kokotekonivelleikkauksen komplikaatiot ovat harvinaisia mutta aiheuttavat merkittävää haittaa yksittäisille potilaille sekä merkittäviä taloudellisia kustannuksia terveydenhuollolle. Tutkimustyö lonkan kokotekonivelleikkauksien tuloksien parantamiseksi on haastavaa, koska komplikaatioiden ilmenemiseen voi kulua jopa vuosikymmeniä ja silloinkin harvinaisten komplikaatioiden todentamiseksi tarvittaisiin suuria potilasjoukkoja. Radiostereometrisellä analyysillä (RSA) voidaan tietyissä tapauksissa kuitenkin ennakoida uusintaleikkauksen riskiä jo kahden vuoden seuranta-ajalla. Tämän väitöskirjan tavoitteena oli laajentaa nykyistä tietoa RSA:lla mitattavan mikroliikkeen syistä ja merkityksestä sekä kehittää RSA-menetelmää lonkan kokotekonivel- ja lonkkamurtumapotilailla.

Ensimmäisessä osatyössä tutkittiin, onko luuntiheydellä merkitystä lonkan sementittömän kokotekonivelen kuppiosan RSA:lla mitattuun mikroliikkeeseen nivelrikkoa sairastaneilla naispotilailla. Toisessa osatyössä tutkittiin kolmiulotteiseen mallinnukseen perustuvan RSA-menetelmän (MBRSA) soveltuvuutta sementittömän lonkan tekonivelen varren tutkimukseen. MBRSA menetelmää tutkittiin ensin fantomia käyttäen ja myöhemmin tulokset varmistettiin lonkan kokotekoniveltutkimukseen osallistuneilla potilailla. Kolmannessa osatyössä selvitettiin kuormituksen aiheuttaman RSA-mikroliikkeen (DLRSA) käyttöä reisiluun kaulan murtumien tutkimuksessa. Neljännessä osatyössä selvitettiin, hyötyisivätkö RSA-tutkimukset moniulotteisesta tilastollisesta menetelmästä.

Lonkan sementittömän kokotekonivelen kuppiosan varhainen mikroliike oli suurentunutta potilailla, joiden luuntiheys oli alentunut. MBRSA menetelmä soveltuu tutkitun tekonivelen varren seurantaan ja käyttöön tulevissa tutkimuksissa. Kehitettyä DLRSA-menetelmää voidaan käyttää reisiluun kaulan murtumien tutkimuksessa. Kolmiulotteisella tilastollisella mallintamisella voidaan havaita yksiulotteisia menetelmiä herkemmin ja spesifisemmin eroja RSA-mikroliikkeessä.

AVAINSANAT: Radiostereometrinen analyysi, lonkan kokotekonivelleikkaus, lonkkamurtuma, osteoporoosi, kolmiulotteinen mikroliike

Table of Contents

Abbreviations

2D	two-dimensional
3D	three-dimensional
ABG	Anatomic Benoist Girard
ANOVA	Analysis of variance
BMD	Bone mineral density
BMI	Body mass index
CAD	Computer-aided design
CI	Confidence interval
CN	Condition number
CT	Computed tomography
CTMA	CT-based implant motion analysis
DLRSA	Differentially-loaded radiostereometric analysis
DLT	Direct linear transformation
DRR	Digitally reconstructed radiograph
DXA	Dual-energy X-ray absorptiometry
EBRA	Einzel Bild Röntgen Analyse
EGS	Elementary geometrical shape
FNS	Femoral neck shortening
HHS	Harris hip score
ICC	Intra-class correlation coefficient
IBRSA	Image-based radiostereometric analysis
ISO	International Organization For Standardization
LMM	Linear mixed-effects model
LOA	Limits of agreement
MANOVA	Multivariate analysis of variance
MBRSA	Model-based radiostereometric analysis
ME	Mean error of rigid-body fitting
MTPM	Maximum total point motion
RE	Reverse-engineered
RSA	Radiostereometric analysis
SD	Standard deviation

TT	Total translation
TR	Total rotation
THA	Total hip arthroplasty
WOMAC	Western Ontario and McMaster Universities Osteoarthritis index

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Finnilä S, Moritz N, Svedström E, Alm JJ, Aro HT. Increased migration of uncemented acetabular cups in female total hip arthroplasty patients with low systemic bone mineral density. A 2-year RSA and 8-year radiographic follow-up study of 34 patients. *Acta Orthopaedica,* 2016; 87(1):48–54.

II Nazari-Farsani S, Finnilä S, Moritz N, Mattila K, Alm JJ, Aro HT. Is model-based radiostereometric analysis suitable for clinical trials of a cementless tapered wedge femoral stem? *Clinical Orthopaedics and Related Rese*arch, 2016; 474:2246–2253.

III Finnilä S, Moritz N, Strandberg N, Alm JJ, Aro HT. Radiostereometric analysis of the initial stability of internally fixed femoral neck fractures under differential loading. *Journal of Orthopaedic Research*, 2019; 37(1):239–247.

IV Finnilä S, Löyttyniemi E, Aro HT. A randomized denosumab trial of total hip arthroplasty – multivariate reanalysis of cementless femoral component 3D migration. Submitted for publication.

1 Introduction

Total hip arthroplasty (THA) is recognized as one of the greatest successes of modern orthopaedics (Learmonth et al. 2007). THA is a treatment-of-choice for advanced hip osteoarthritis – a disease that is one of the most common and, also, most debilitating conditions affecting the elderly worldwide (Murphy et al. 2010, Glyn-Jones et al. 2015). In Finland alone over 10 000 primary THAs are performed each year – osteoarthritis, femoral neck fracture, and avascular necrosis being the three most common indications for the procedure (Finnish Arthroplasty Register 2021). The worldwide incidence of THA has been estimated to exceed two million operations a year by 2030 (Pivec et al. 2012). However, revision surgery as a result of failing primary joint arthroplasty is associated with significant economic and humane burden (Weber et al. 2018). Consequently, there is a tremendous incentive to improve and innovate on the initial success of THA.

The research focused on improving the initial success of THA is not straightforward. The clinical manifestation of complications can take decades (Michelson and Riley 1989). To solve this issue, radiostereometric analysis (RSA) has emerged as a key research method in the clinical introduction of THA components (Pijls and Nelissen 2016). RSA is the most accurate method for the determination of early implant micromotion in relation to host bone *in vivo* (Kärrholm et al. 2006). Using RSA, the 10-year revision rates of THA, related to aseptic loosening, can be predicted utilizing data from only 15 to 25 patients at two years postoperatively (Valstar et al. 2005, Pijls et al. 2012a, de Vries et al. 2014, van der Voort et al. 2015). Therefore, RSA has a pivotal role in enabling current and future innovation related to THA.

The implant-related determinants of early micromotion have understandably received considerable focus in RSA studies. However, many of the patient-related factors for implant stability have not been systematically studied with RSA. This is despite the fact that implant fixation is fundamentally an interplay between the implant and host tissues (Sundfeldt et al. 2006). Indeed, optimizing patient selection and, possibly, introducing targeted medical interventions may be important factors for achieving an improved outcome for THA. Also, in RSA studies focused on

optimizing implant design, controlling for the patient-related variation is a valuable asset for maximizing the proportion of implant-related variation in the RSA data.

Furthermore, the RSA methodology is still undergoing continuous development. The introduction of model-based RSA (MBRSA) has enabled the study of unmodified clinical orthopaedic implants (Valstar et al. 2001, Kaptein et al. 2003). This has considerable advantages over manufacturing, and acquiring regulatory approval for, custom RSA-marked implants. However, when using model-based RSA, the shape of the implant and the quality of the implant surface model may affect the accuracy and precision of the method (Valstar et al. 2001). Therefore, validating the method is mandatory for each individual implant type in preparation for a clinical study (Kaptein et al. 2006).

In Finland, femoral neck fracture has been the second most common indication for THA after osteoarthritis, in recent years (Finnish Arthroplasty Register 2021). A joint preserving internal fixation of a femoral neck fracture may be more beneficial in select patient populations. Still, optimizing patient selection and the surgical intervention remains a challenge (Probe and Ward 2006). Meanwhile, the clinical study of fracture healing has been hampered by the lack of accurate research methods (Corrales et al. 2008, Augat et al. 2014, Morshed 2014, Bizzoca et al. 2020). Differentially-loaded RSA (DLRSA) has recently emerged as a potential research method to overcome this issue (Chehade et al. 2009), but the method has not yet been validated for femoral neck fractures.

Steps toward the standardization of the RSA have been undertaken in the past two decades (Valstar et al. 2005). This is critical to maintain the comparability of the results between multiple centres and researchers. However, these standardization efforts did not cover the statistical analysis of the complex three-dimensional (3D) micromotion data produced by RSA. The current common practice among RSA studies seems to be to analyse the data axis-by-axis using univariate statistical methods. A comprehensive multivariate statistical model could provide both improved statistical power and reduced risk of false-positive results but, previously, their application to RSA data analysis has not been studied systematically.

This doctoral thesis was aimed at improving the present understanding of the RSA-measured micromotion and refining the related RSA methodology so as to facilitate future clinical research of total hip arthroplasty and hip fracture patients. The effect of quantitative systemic bone mineral density (BMD) on RSA-measured micromotion was studied for a cementless acetabular cup using marker-based RSA technique. The applicability of MBRSA for the study of a tapered-wedge cementless femoral stem was confirmed using both a phantom model and clinical RSA data. DLRSA was validated for the study of femoral neck factures in a clinical patient cohort. Finally, a multivariate statistical method was compared to univariate methods in the analysis of clinical and simulated RSA data.

2 Review of the Literature

2.1 A Brief History of RSA

Stereoscopy alludes to the perception or illusion of depth created by the binocular vision system. From minute changes in the perspective of each eye – parallax – the relative distance of an object can be estimated by the visual cortex of the brain (Howard and Rogers 2008). Taking advantage of the binocular system, stereographs can be used to create the illusion of depth by showing each ocular a two-dimensional photograph with a slightly different viewpoint. A special device called stereoscope can also be used to help view stereographs (Wheatstone 1838). This technique is not limited to photographs, however, as stereoscopic radiographs can be obtained by using two separate x-ray source placements, in other words, foci.

Indeed, before the emergence of computed tomography (CT) and computer-generated 3D reconstructions, stereoscopy was the first application of stereoscopic radiographs that saw more widespread adoption. By obtaining two radiographs with different foci of an object, stereoscopy could be utilized to provide a perception of relative depth in radiographs (Davidson 1916). Stereoscopic radiographs grew in popularity in the early 20th century and were often routinely used in the clinical setting. However, while stereoscopy proved a useful technique in creating the perception of depth, it provided little quantitative advantage over plain radiographs: only relative depth could readily be judged from the stereographs. Taking multiple radiographs also resulted in additional radiation exposure. Therefore, the technique later saw a steep decrease in its use, as the effects of radiation on the human body were better understood (Curry et al. 1990). More recently, CT has superseded stereoscopy in clinical diagnostics and this form of radiographic stereoscopy has been all but forgotten.

The first attempts at radiostereometry, i.e. making quantitative measurements from stereoscopic radiographs, date back to the very early days of radiography at the end of the 19th century. As a crude form of radiostereometry, Stanley Davidson (1898) used a custom measurement device – Davidson's localizer – to reconstruct the spatial coordinates of features in stereoscopic radiographs (**Figure 1**). The device was first used to expose a radiographic film twice from distinct known positions with the object kept as still as possible between the exposures. The Davidson's localizer

then used silk threads to represent the path of x-rays from the two foci to the features on each corresponding exposure. The intersection of these two threads represented where the measured feature had been in space during the exposures and the 3D coordinates of this point could then be measured.

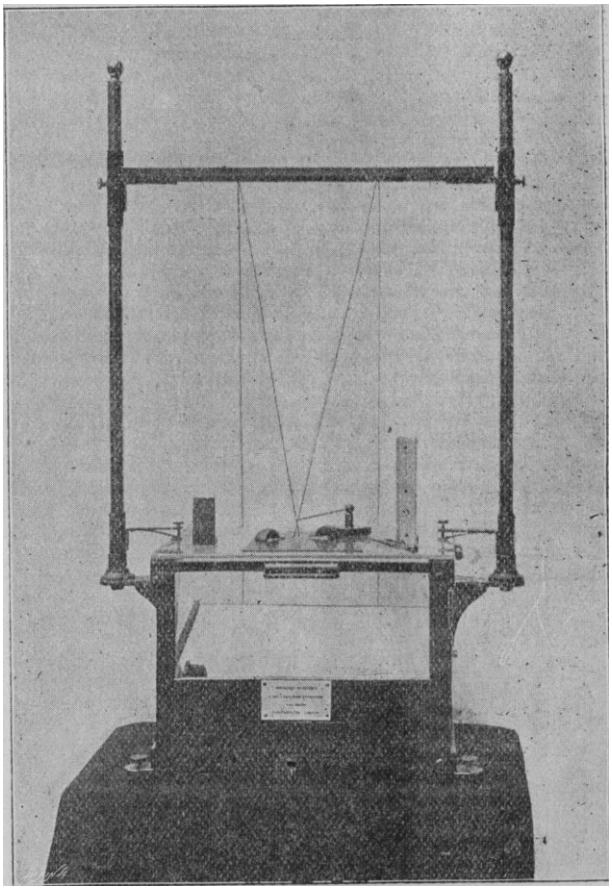


Figure 1. Davidson's localizer with threads projecting from the two distinct exposure positions of the x-ray focus to the radiographic film. Reproduced with the publisher's permission from the original work by Stanley Davidson (1898).

Davidson's method used a single radiographic film, with one x-ray source moved between the exposures – very little in the way of special instrumentation was required. One of the proposed uses was the precise localization of foreign objects within patients' bodies. In the 1930s, mathematician Margherita Piazzolla Beloch introduced a similar, if slightly improved, method where two separate x-ray sources were used to simultaneously expose two separate radiographic films at a right angle

to each other (Piazzolla Beloch 1936). Akin to Davidson's method, a rigid fixture was used to obtain the radiographs and the fixture was later used to reconstruct the path of the x-rays with the help of threads. Yet, little came of either of these methods as evidenced by the lack of related literature later in the 20th century (Selvik 1989).

Although measurements on the features of radiographs could be made with either Davidson's or Piazzolla Beloch's method, the coordinates of a single bony feature or foreign body are often not, in and of themselves, useful. The relevant question the researcher many times finds him- or herself pondering is: have the structures of interest moved in relation to each other between two timepoints and, if so, how much and to which direction? Questions related to movement over time, i.e. kinematics, can only be answered if first a reliable reference coordinate system can be established. A related breakthrough was made with the introduction of metallic implantable pins in 1950s for the measurement of facial bone growth (Björk 1968). Of the available materials for bone markers, tantalum proved the most promising. Tantalum has both high elemental density and atomic number – both important physical properties for blocking energetic electromagnetic radiation such as x-rays – which made it easily visualizable on radiographs even when implanted in relatively radiopaque tissue, such as bone (Björk 1968, Kärrholm et al. 2006). Additionally, elemental tantalum proved to be biologically inert and, yet, able to osseointegrate with the host bone (Aronson et al. 1985). Indeed, tantalum markers made for a highly ideal implantable reference coordinate system within the human body, especially bone tissue, facilitating the radiographic study of bone kinematics.

During the late 19th and early 20th century the theoretical foundations for analytical photogrammetry were laid. These scientific advances would later prove crucial for the development of RSA. Of special note are the equations for projections between planes that were derived by Otto Von Gruber in 1924 (Doyle 1964). Swedish scientist Bertil Hallert eventually applied these principles to radiography (Hallert 1954, 1970). Hallert showed that by utilizing a known 3D grid of radiographic markers – later termed as "calibration cage" – the precise configuration of the radiographic film and focus could be reconstructed mathematically. Conversely, this meant that the path of the x-rays from the radiographic focus to the film could be precisely reconstructed. By combining this concept with stereoradiography, Hallert replaced Davidson's silk threads with far more robust mathematical equations. An added benefit of Hallert's method was that the stereometric measurements were accomplished without prior knowledge as to the precise positioning of either of the radiographic films or the x-ray foci. This eliminated considerable sources of error. Astonishing measurement precision of 10 to 50 µm could be achieved with Hallert's method (Hollender 1964).

The modern era of radiostereometry began in the 1970s with the pioneering work of Göran Selvik at Lund, Sweden (Selvik 1989). Selvik successfully built on the

groundwork laid by Hallert: He utilized tantalum as reliable implantable radio-opaque markers and, finally, introduced rigid body kinematics into the process. That is, he used groups of three or more tantalum markers to define rigid bodies. A rigid body could then be used to define a reference coordinate system in relation to which the motion of other rigid bodies could be calculated and reported (Selvik 1989). The final obstacle of the computationally intensive optimization problems involved with rigid body kinematics was also overcome with the aid of computer-based algorithms (Selvik 1989). Selvik called his method roentgen stereophotogrammetric analysis. His method enabled the precise study of micromotion within the human bone tissue for the first time creating a whole new field of orthopaedic research (Kärrholm et al. 2006).

Early on, in the field of orthopaedics, RSA was used successfully to study the physiology of bone tissue – i.e., growth, fracture healing and joint kinematics (Kärrholm 1989). Later, with the initial successes of total hip and knee arthroplasty in the 1970s and 1980s the physiology and pathophysiology of orthopaedic implants within the human body became of great interest (Kärrholm 1989, Kärrholm et al. 2006). As competing methods lacked the necessary precision, RSA emerged as the method of choice for studying early migration of hip and knee implants in vivo (Kärrholm 1989, Malchau et al. 1995). Indeed, in his posthumously published review article on RSA, Selvik himself considered the study of total hip arthroplasty "-- the most successful application of RSA" (Selvik 1990).

It could be argued that the clinical value of RSA in the study orthopaedic implants turned out even greater than what Selvik and colleagues had expected. The reason for this is twofold. First, early in the study of hip and knee implants, evidence began mounting supporting the hypothesis that early migration was linked to implant loosening and need for revision years later (Ryd 1992, Kärrholm et al. 1994). In other words, RSA seemed to have predictive power of clinical failure rates due to aseptic loosening years ahead of time, which has recently been confirmed in meta-analysis –level studies (Pijls et al. 2012c, 2012a, de Vries et al. 2014, van der Voort et al. 2015). Second, the high precision of RSA and a surprisingly low population-level variance in the early migration allowed for very small sample sizes of only 15 to 25 patients to arrive at satisfactory estimates of the population-level implant failure rates (Valstar et al. 2005).

The early successes of RSA prompted growing international interest and multiple centers have since started contributing to the growing body of RSA research data. In an attempt to maintain comparability of data between research centers, the first steps towards the standardization of the RSA method were taken a little more than a decade ago (Valstar et al. 2005). Also, an ISO-standard was issued for RSA (International Organization For Standardization 2013). The guidelines seemed to

improve the methodological quality of subsequent studies, but still left room for improvement (Madanat et al. 2014).

The advances made in the development and standardization of RSA have coincided with considerable issues raised in the development of orthopaedic implants. In the recent decades new implant designs have been introduced to clinical use with practically no evidence supporting their deployment – sometimes with catastrophic results. Compared to existing designs implants with two- to tenfold greater risk of revision have reached the market (Nieuwenhuijse et al. 2014, Pijls and Nelissen 2016). These mistakes have often only become evident after years of clinical use from the excessive number of revisions. By the time the high failure-rates have become evident, hundreds or even thousands of patients have sometimes been exposed to these implants with tremendous humane and financial consequences (Ohlin 1990, Malchau et al. 1993, Muirhead-Allwood 1998, Norton et al. 2002, Hauptfleisch et al. 2006, Gilbert et al. 2009, Sedrakyan 2012).

To prevent the introduction of flawed implant designs into wider clinical adoption in the future, phased introduction of new implants has been proposed akin to how new medications are phased into clinical use to minimize potential risks (Nelissen et al. 2011, Pijls and Nelissen 2016). Due to its unique predictive power of long-term survivorship from relatively small sample sizes and short follow-up times, RSA has been proposed as the first step in the clinical introduction of new implants (Malak et al. 2016, Pijls and Nelissen 2016). Indeed, it seems that RSA has secured its place as an invaluable orthopaedic research tool for the foreseeable future.

2.2 RSA Methodology

2.2.1 Marker-Based RSA

The basic principles of RSA have remained remarkably unchanged since their conception by Göran Selvik despite some algorithmic improvements and amendments (Probst et al. 1978, Selvik et al. 1983, Selvik 1989, 1990, Valstar et al. 2001, Börlin et al. 2002, Kaptein et al. 2006, Kärrholm et al. 2006). The principal phases of a radiostereometric analysis can be divided into four distinct parts: implementation of markers (and possible implants), radiographic examination, measurement of the radiographs and mathematic calculations (Kärrholm 1989).

Implementation of Markers

The marker-based RSA, as described by Selvik, relies mostly on analyzing the motion of radiopaque markers within the study subject (Selvik 1989). At present, spherical tantalum beads are the gold standard for bone markers among RSA studies

(Valstar et al. 2005). Special instrumentation has also been developed to aid in implementation of markers (Aronson et al. 1974). In the simplest case of studying a single degree of freedom, implanting just two tantalum markers can suffice for making a meaningful measurement. The difference in distance between these two points can be calculated between a principal examination and any number of follow-ups. This methodology can be utilized in the study of bone growth, for example (Selvik 1990).

However, to estimate and represent the motion of implants or segments of bone in a meaningful manner, it is necessary to introduce the concept of rigid body kinematics: Rigid bodies are, by their theoretical definition, solid objects whose internal distances do not change over time or by application of external forces. In the context of an RSA study, a segment of bone or an implant can be considered a rigid body. By attaching at least three RSA markers to a rigid body, we can measure its 3D pose (orientation and position) in a laboratory coordinate system with RSA (**Box 1**). With one rigid body acting as a reference coordinate system within the body, the relative position of any number of additional markers or rigid bodies can then be calculated in relation to this reference coordinate system. Finally, relative motion of these rigid bodies (i.e., kinematics) can be estimated between two or more examinations as changes in their relative poses. This relative motion between, for example, a bony reference coordinate system and a rigid body formed by an implant is the principal measurement of interest in most implant studies and what is conceived as the implant migration in RSA studies.

BOX 1. Minimum requirements for the pose determination of a rigid body

A 3D rigid body has six degrees of freedom in motion (Selvik 1989). That is, the changes in position and orientation of said rigid body can be described by a minimum of six independent parameters. Therefore, a minimum of three distinct 3D measurement points (i.e., RSA markers) on a rigid body are necessary to definitively determine its motion. Proof follows:

A measurement of motion at a single 3D point on a rigid body provides only three degrees of measurement (independent parameters) which is mathematically insufficient to determine the six degrees of freedom of rigid body motion. A second 3D point is somewhat counterintuitively not sufficient to solve all six degrees of freedom. This is because the interpoint distances between different parts of a rigid body are, per the definition of a rigid body, constant. Thus, a second 3D measurement point adds only two degrees of measurement as the interpoint distance between the first and second measurement point is constant. A third measurement point adds only one degree of measurement, the sixth, as two interpoint distances are already known for it. Subsequent measurement points do not add degrees of measurement but may aid in the precision and accuracy of movement estimation by way of signal averaging (Tagare 1993).

The configuration of the markers on a rigid body has a great impact on the precision of pose estimation. A good spatial distribution of markers is necessary for the precise determination of rotations. In the extreme case of collinear markers, the

determination of rotation around the line formed by the markers is impossible (Söderkvist and Wedin 1993). In studies involving small bones, joints or implants, an ideal scatter of RSA markers is sometimes impossible to achieve due to physical constraints on the marker scatter. In these cases, concessions in the measurement accuracy are sometimes necessary (Valstar et al. 2005).

The stability of markers in relation to surrounding bone or an implant is imperative for the precision of the method. In studies dealing with large radiodense orthopaedic implants, markers may also be obscured in radiographs by the implant itself. Therefore, great care should be taken in placing the markers to maintain their stability and visibility. Inevitably, individual markers may become loose or impossible to visualize in some, or all, of the stereoradiographs. Consequently, it has been recommended that 6 to 9 markers should be used for determination of a bony rigid body (Valstar et al. 2005). Additional markers also help overcome noise in the measurement system through what is effectively static-signal averaging (Tagare 1993). Marking implants is oftentimes expensive and time consuming which is why only 3 implant-markers can be considered acceptable (Kaptein et al. 2003, Valstar et al. 2005).

Radiographic Examination

An RSA imaging setup requires two separate foci to obtain stereoradiographs. The foci are placed such that the paths of the x-rays converge on the objects under study. In modern setups, each focus exposes a separate radiograph. Two radiographs of the object under study are thereby obtained, each from the viewpoint of the corresponding focus. Typically, two separate x-ray sources are utilized to expose the radiographs simultaneously. This ensures that the object under study cannot move between exposures which could in turn compromise the measurement precision (Selvik 1990).

The precision of RSA hinges on determining the precise geometry of the radiographic setup (**Figure 2**). This is enabled through highly accurate calibration cages. These calibration cages consist of RSA markers in known relative positions and therefore allow the mathematical reconstruction of the radiographic setup. Each radiographic setup used for RSA must be calibrated by obtaining a stereoradiograph of a calibration cage at least once. Best precision is achieved, however, if every pair of radiographs is exposed through a calibration cage. This eliminates the effects of changes in the geometry of the radiographic setup between examinations (Selvik 1990).

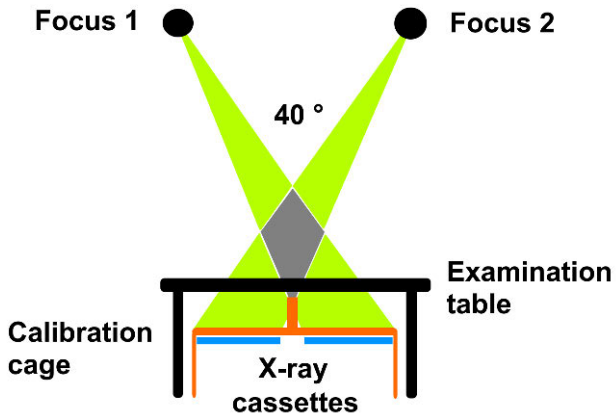


Figure 2. The schematic radiographic setup for a typical uniplanar RSA examination.

Based on the type of the calibration cage, the radiographic configuration used in RSA can be divided into two major categories: uniplanar and biplanar. In the uniplanar technique, so-called fiducial markers are on a single plane for both radiographs. In the biplanar technique, the setup is otherwise similar, but the fiducial markers of each radiograph are on two different planes. In the biplanar technique the planes are typically at right angle to each other (Selvik 1989, Valstar et al. 2005). In a conventional biplanar calibration cage the calibration markers surround the object under study. Thus, smaller joints in the extremities are amenable to study with a biplanar cage. A uniplanar cage can more readily be situated under an examination table, for example, allowing larger structures such as the hip joint or the spine to be imaged (Kärrholm 1989).

Important considerations relate to the type of the calibration cage and the precision of the RSA method. A typical uniplanar cage has unsymmetrical precision in the so-called in-plane and out-of-plane axes of motion. This is due to two principal reasons: First, the typical uniplanar setup utilizes a less than 90 degrees of convergence for the two x-ray beams (**Figure 2**) which means that motion in the out-of-plane axis results in a relatively smaller movement of the corresponding radiographic projection (Cai et al. 2008). Secondly, a uniplanar cage uses extrapolation from the control markers to perform measurements outside the calibration cage in the direction of the out-of-plane axis. Thus, the precision of measurements gets progressively worse as the distance from the calibration cage increases (Choo and Oxland 2003). The difference in precision between different axes should be considered when analyzing the results obtained using a uniplanar calibration cage.

RSA is a demanding technique and training the relevant staff for obtaining good quality RSA radiographs is necessary (Muharemovic et al. 2018). Also, although the

patient can move in relation to the calibration cage between examinations, the patient should be imaged in an anatomical reference position at least once. This is done to orient the reference coordinate system in the patient in a repeatable manner between patients (Valstar et al. 2001). Alternatively, features on the stereoradiograph can be used to orient the reference coordinate system (Laende et al. 2009). Whatever the method for orienting the reference coordinate system it is crucial to standardize the orientation between studies of similar implants to maintain comparability between RSA studies (Valstar et al. 2005).

Measurement of Radiographs

Locating both stereographical projections of each RSA marker as accurately and precisely as possible underlies any subsequent attempts at reconstructing the 3D positions of said markers. In the 1990s and early 2000s, digital methods for the measurement of RSA radiographs superseded the previously used manual method (Börlin et al. 2002). Computerized algorithms enabled the use of more data from the radiographs, which resulted in even better precision than previously had been possible. These algorithms also made the process of marker detection semi-automated which resulted in improvements of interobserver agreement. Furthermore, the tedious marker-by-marker measurement and marker-matching was made slightly faster with the help of automated computer algorithms. Also, the related progression to digital imaging plates has eliminated potential error sources of digitizing film-based radiographs (Bragdon et al. 2004, Mäkinen et al. 2005).

Mathematic Reconstruction of the RSA Setup and Migration

The mathematic calculations in an RSA study, involving rigid bodies, can be divided into two major categories (Selvik 1989): First, resolving the geometry of the radiographic setup and reconstructing the 3D positions of so-called object markers. Second, resolving the rigid body transformations (i.e., changes in rigid body position and orientation) between examinations. In an idealized world, geometry and algebra could be used to arrive at exact solutions to both of these mathematical problems. However, the measurements made on the radiographs are not arbitrarily exact. Additionally, the rigid bodies formed by the RSA markers are never strictly rigid as RSA markers may become loose and the underlying bone structures in the human body undergo constant remodeling. For this reason, arriving at an exact solution for either of the posed mathematical problems is not practically possible (Selvik 1989, Choo and Oxland 2003). This issue can be circumvented by approximating an optimal solution with the aid of computers (**Box 2**).

BOX 2. Computational geometry and optimization algorithms

Problems where arriving at an exact solution is impossible or computationally infeasible are common in computational geometry (Ausiello et al. 1999). Thus, approximating a solution is necessary. In the case of nearly intersecting lines, such as reconstructing the 3D position of an RSA marker, an optimal approximation can be found algebraically (Selvik 1989). On the other hand, approximating rigid body transformation represents a non-linear least-squares problem where arriving even at an optimal solution is implausible. However, a sufficiently accurate approximation of the relevant rotation and translation parameters can be achieved by minimizing a cost function through an algorithm such as Gauss-Newton (Selvik 1989). The estimation of the rigid body transformations involves iterative calculations on complex least-squares equations. Solving them without the aid of computers would be prohibitively time-consuming to do. This is one of the reasons why RSA has greatly benefitted from the evolution of computer technology.

The calibration cage plays a central role in approximating the geometry of the radiographic setup. According to the methodology described by Selvik (1989), each calibration cage contains fiducial markers and control markers. The fiducial markers in a calibration cage (**Figure 3**) are placed on a plane in front of each radiograph (Selvik 1989). At least four fiducial markers are then used for the calibration of each radiograph through projective transformation as described by Hallert (1970). The obtained calibration values are used to resolve any subsequent marker projection on the film into a 2D *laboratory coordinate system.* This laboratory coordinate system is therefore affixed to the calibration cage on the plane of the fiducial markers. The control markers which are placed at a known distance from the fiducial plane are subsequently used to calculate projection lines from the foci to the radiographic film. The point at which these projection lines are closest to intersecting are used as the estimated position of each corresponding focus. Knowing the position of the foci, the position of any RSA marker visible on both radiographs, in relation to the laboratory coordinate system, can readily be calculated as the position where both projection lines of the marker intersect or, in fact, are closest to intersecting (**Box 2**).

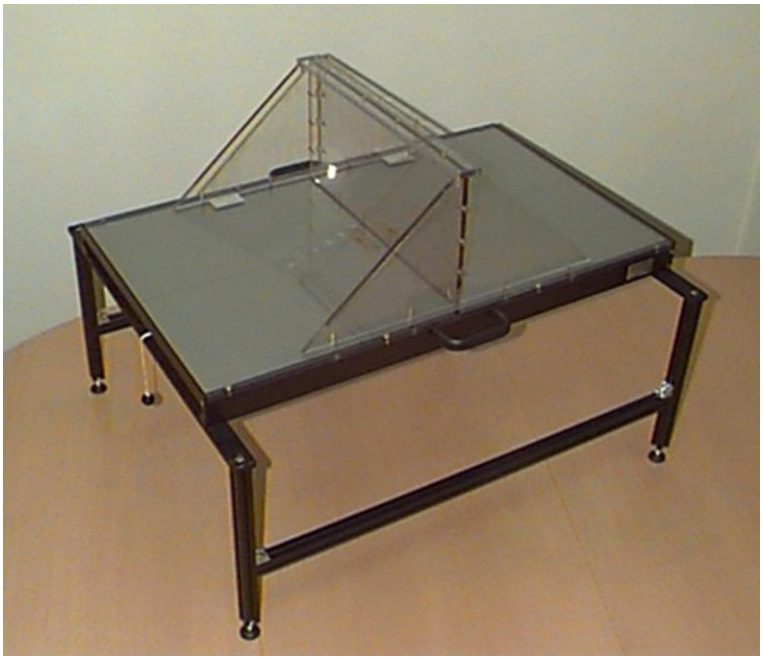


Figure 3. A uniplanar RSA calibration cage with RSA markers inserted in precisely defined positions within the sheets of transparent plastic. Fiducial markers are placed on the flat surface above the X-ray cassettes. Control markers are placed on top of the vertical platform. The same control markers are projected onto each radiograph.

Since the early days of RSA projective transformation has been superseded by Direct Linear Transformation (DLT) (Selvik 1990). With DLT all fiducial and control markers can be used for the approximation of necessary calibration values in a single step streamlining the related computations. In fact, DLT does away with the need for separate fiducial and control markers as long as the calibration markers have sufficient spatial distribution. Also, DLT has been shown to provide at least equal or in some cases greater accuracy compared to the original procedure described by Selvik (Choo and Oxland 2003). A related benefit of DLT is that the calibration markers need not be confined to planes as long as their 3D position is known (Choo and Oxland 2003). This offers greater potential for flexibility as to the configuration of the RSA setup.

After the geometry of the radiographic setup has been established, the concepts of rigid body kinematics can be implemented: First, at least two stereoradiographs of the rigid bodies under study are obtained. For each stereoradiograph, the 3D positions of all object markers (i.e., markers attached to a rigid body) are calculated as described above. Then, markers forming the reference coordinate system within the study subject are identified and the transformation of a reference rigid body between the two stereoradiographs is calculated. This is done by obtaining 3D

migration parameters that minimize the root mean square value of difference in the positions of all markers belonging to the reference rigid body between the stereoradiographs (**Box 2**). Next, all measured RSA marker positions in the latter RSA examination are transformed according to the observed migration of the reference rigid body. Thus, the displacement of the reference coordinate system (i.e., study subject) in relation to the laboratory coordinate system between the examinations is eliminated. Finally, using the transformed coordinates, we can estimate the displacement of additional rigid bodies (e.g. implants) in relation to the reference rigid body by simply repeating the same process for each corresponding rigid body (Selvik 1989).

2.2.2 Model-Based RSA (MBRSA)

Conventional RSA relies on markers being attached to an implant. Marking the implants and gaining regulatory approval for these modified implants can be time-consuming and cause a considerable drain on research funding (Valstar et al. 2001, Kaptein et al. 2003). Additionally, RSA of radiodense orthopaedic implants runs the risk of obscuring markers attached to both the implant and the surrounding bone (Valstar et al. 2001, Kaptein et al. 2006). In some cases poor marker-visualization can lead to either unnecessary radiation exposure due to repetitive imaging or patients being excluded from the analysis entirely (Muharemovic et al. 2018). It has even been stated that difficulties in marker visualization may have considerably hindered RSA studies of some implant types (Valstar et al. 2001). Thus, there is a demand for alternative methods of measuring implant motion in RSA studies.

Elementary Geometrical Shape Modelling

Early in the development of RSA, in addition to markers, well-defined features of an implant were utilized as measurement points. These included the femoral head of an implant and an indicator ring on an acetabular implant (Baldursson et al. 1979, Mogensen et al. 1982). For both of these objects, the center coordinate could readily be determined from their two-dimensional (2D) radiographic projections. Thereby they essentially served as additional RSA markers. Later, more complex geometrical shapes have been analyzed. For example, conical and cylindrical shapes in hip stems have been utilized to provide measurements of a line passing through the implant. The opening of an acetabular cup has also been used to provide a partial estimate of the cup's orientation (Valstar et al. 1997). By combining several of the aforementioned measurement points it is possible to perform RSA without attaching any implant-markers (Kaptein et al. 2006). The concept of using implant features and their related geometrical models for acquiring measurement points has been

termed *Elementary Geometrical Shape modeling RSA*, or EGS-RSA (Kaptein et al. 2006).

EGS modelling revolves around establishing a geometric relationship between the observed 2D projection of an implant and its 3D position or orientation (Kaptein et al. 2006). Such relationships are easier to establish and generalize for parts of implants with rotational symmetry around a line or a point as this reduces the number of parameters that need to be solved for. The obvious limitation of EGS is that only implants with relatively simple and consistent geometric shapes can be measured. This limits the types of implants to which EGS-RSA can be applied (Kaptein et al. 2006). Also, a sufficient number of measurement points with adequate spatial distribution, similar to marker-based RSA, is a prerequisite for the precision of the method (Kaptein et al. 2006).

Implant Surface Model-Based Analysis

Some of the shortcomings of EGS-RSA can be overcome by approaching the model-based RSA concept from an alternative angle. By using a known 3D surface model of an implant, it is possible to reconstruct the shape of the resulting 2D projection in a radiographic setup with known geometry (Valstar et al. 2001). Determining the orientation of an implant becomes, then, an issue of finding a position and orientation for the 3D implant model that results in a projection that best describes the observed radiographic projection of the real-world implant. Mathematically, this is a matter of minimizing the error between the observed and simulated implant projection using the position and orientation of the 3D model as the input parameters. This implant surface model-based concept was proven by Valstar et al. (2001). Subsequently this method has evolved even further and is nowadays considered a valid alternative to marker-based RSA (Kaptein et al. 2003, Trozzi et al. 2008, Hurschler et al. 2009, Seehaus et al. 2009).

However, the quality of the 3D surface models is paramount to the success of the method. Already in the pilot study Valstar et al. (2001) noted that the precision and accuracy of the method were sensitive to errors in the 3D implant model. Also, rotational symmetry degrades the precision of model-based RSA analogous to how collinear markers are deleterious for the precision of marker-based RSA (Valstar et al. 2001). The implant surface models can come from two principal sources. The *computer-aided design* (CAD) model, used for developing and manufacturing the implant, can be used for the analysis (Valstar et al. 2001). Alternatively, a reverse-engineered (RE) model can be acquired by a 3D scanner that measures the implant and creates a surface model for it (Kaptein et al. 2003).

Of the two methods, using CAD models is considered the less precise approach (Kaptein et al. 2003). Refining manufacturing phases like machining, additional

coatings and surface roughening etc. may not be represented by the initial CAD design (Valstar et al. 2001, Kaptein et al. 2003). Reverse-engineering the surface model from the finalized implant removes these error sources and, indeed, RE models have shown superior accuracy and precision compared to CAD models (Kaptein et al. 2003). The best results could theoretically be achieved with using a case-by-case implant-specific RE model (Kaptein et al. 2003). However, obtaining a RE model for each implant is often not feasible due to financial and logistical restrictions. Thus, a single representative RE model or a CAD model is usually employed in the clinical setting. At least in the case of a cementless hip short-stem using a single representative model resulted in no considerable concessions in accuracy or precision (Seehaus et al. 2016).

Even though the implant surface MBRSA has been proven to have good accuracy and precision for many implant types it does have its theoretical weaknesses. If a 3D object has an axis of symmetry, like for example most acetabular cups do, its 2D projections do not change when the object rotates around said axis. This is a recognized shortcoming of MBRSA (Valstar et al. 2001, Kaptein et al. 2003). Even in cases, where there are no axes of prefect symmetry, the object can be symmetrical enough on some axes to degrade the precision of the model-based method: rotation of most femoral stems around their longitudinal axes, for example, causes only minute changes in the 2D projection of the stem degrading the measurement precision of said rotation. For femoral stems, this issue can be somewhat alleviated by adding the femoral head into the model (Prins et al. 2008). Still, rotational symmetry is an issue that must be a consideration in MBRSA studies.

Marker-Free RSA

The surface model-based concept can be extended to completely marker-free RSA. By using bone models obtained with CT the motion of the corresponding bone sections can be reconstructed with MBRSA (Seehaus et al. 2012). This methodology has allowed, for example, joint kinematics to be studied without attaching markers (Hansen et al. 2017). However, the precision reported for the methodology has been somewhat disappointing (Seehaus et al. 2012).

Recently tentative results have been published on an improved marker-free RSA methodology. In conventional MBRSA only a portion of available information in a radiograph is used. Instead of fitting only the virtual contours of a model projection to the obtained radiographs, the volumetric data from a CT scan can be used to simulate the complete radiographic projection including changes in radiodensity (de Bruin et al. 2008). Analogous to the original model-based RSA concept, an optimizing algorithm can then be used to find a pose for the 3D model that produces a *digitally reconstructed radiograph* (DRR) as close to the observed real-world

radiograph as possible. This method has been called *image-based RSA* (IBRSA) to distinguish it from earlier iterations of model-based RSA (de Bruin et al. 2008). IBRSA has demonstrated accuracy comparable to MBRSA and in a direct comparison has showed superior precision to conventional MBRSA in marker-free analysis (de Bruin et al. 2008, Hansen et al. 2018). IBRSA is still in its infancy, however, and requires further research to confirm the promising preliminary reports.

2.2.3 Inducible Micromotion under Differential Loading

Already in the 1980s, it was shown that RSA could be used to detect stress-related or inducible micromotion of orthopaedic implants (Mjöberg et al. 1984, Ryd 1986). In other words, the instant micromotion of implants due to external forces was studied using RSA. The method of studying RSA-measured micromotion under quantified external loads has later been termed *differentially loaded RSA* (DLRSA) (Chehade et al. 2009). Measuring inducible micromotion with RSA provides an opportunity to observe the instantaneous biomechanical properties of the bone-implant or bone-bone interface *in vivo* (Ryd 1986, Chehade et al. 2009). A mechanical disconnect between an implant and host bone, as demonstrated by excessive inducible RSA micromotion, has been viewed as evidence for the lack of osseointegration – a phenomenon that remains extremely difficult to quantify *in vivo*(Ryd 1986, Ryd et al. 1990). This is in contrast with traditional RSA where the detection of static implant or fracture displacement likely provides only indirect evidence of mechanical or biological processes that have already occurred previously (Sundfeldt et al. 2006).

For orthopaedic implants, inducible micromotion has been utilized in the study on tibial components in either unicompartmental or total knee arthroplasty by multiple authors (e.g., Ryd 1986, Petersen et al. 1999, Regnér et al. 2000, Hansson et al. 2005, Wilson et al. 2010, Augat et al. 2014, Lam Tin Cheung et al. 2018, Laende et al. 2019a, 2019b). It has also been used for the study of hip stems (Glyn-Jones et al. 2006, Ferguson et al. 2018), acetabular cups (Digas et al. 2013) and ankle implants (Dunbar et al. 2012). Still, the clinical causalities of inducible micromotion have not been completely established for any implant type. The significance of DLRSA for the study of fractures will be discussed in more detail later in the thesis.

The aim of DLRSA is to perform quantitative measurement of the biomechanics either at a fracture site or a bone-implant interface. A major related issue is standardizing the load that the studied site is subjected to. Voluntary motion and force generation by patients is notoriously unreproducible and can potentially overwhelm the accuracy of RSA measurements with biological variability leading to nonsensical results (Ryd et al. 2000). Carefully quantifying the force used to generate a specific displacement enables the direct estimation of mechanical

properties, i.e. stiffness, at the fracture site or bone-implant interface (Chehade et al. 2009). Quantifying stiffness is of special interest in the study of bone fracture healing (Chehade et al. 1997, 2009, Claes and Cunningham 2009).

2.2.4 Accuracy, Precision, and Power Analysis

Accuracy and Precision

In the context of RSA, a clear understanding and agreement on the concepts of accuracy and precision are necessary. As the RSA methodology has been consolidated as part of the ISO standards it is reasonable to adopt the ISO definition of both precision and accuracy (International Organization For Standardization 1994): The ISO standard defines accuracy of a measurement method as the combination of its trueness and precision. Trueness, on the other hand, is the average measurement's closeness to a true value. Thus, trueness is the opposite of bias. Precision is defined as the agreement between multiple test results – i.e., reproducibility. Accuracy thereby encompasses both systematic and random errors in a measurement system.

Trueness of a measurement system can only be evaluated if a superior measurement method or estimate of the true value exists. It is understandable then that, besides zero motion evaluations, evaluating the trueness – and consequently accuracy – of an RSA setup is infeasible in the clinical setting as there is no reference standard with which the RSA results can be compared (Valstar et al. 2005). Therefore, *in vitro* phantom studies, in which the motion of rigid bodies can be accurately controlled, need to be conducted instead (Valstar et al. 2005). A phantom study should be conducted for each new RSA setup to validate its accuracy and precision before proceeding to clinical studies (Valstar et al. 2005).

While accuracy of RSA cannot readily be determined in the clinical setting we can still estimate it indirectly. Systematic error is typically negligible compared to random error in RSA examinations. As a consequence, random error, which can be evaluated through double examinations, is the chief determinant of accuracy in a typical RSA setup (Selvik 1989). Thus, by evaluating clinical precision we can arrive at a reasonable estimate of the clinical accuracy in an RSA study. Therefore, assessing the clinical precision of RSA is a crucial part of all RSA studies (Valstar et al. 2005). The precision of a measurement system can be estimated through repeat examinations (Ranstam et al. 2000). In the context of RSA this is achieved by performing double examinations on a patient within short time intervals (Valstar et al. 2005). For the calculation of precision, zero true motion between these double examinations is assumed. Yet, random measurement errors will result in observed motion between the double examinations. By repeating the double examination

process on multiple patients and follow-ups the mean level of random measurement error in the study setup over the study period can be estimated. Using these estimates, confidence intervals for containing a double measurement, e.g. precision of the study setup, can be calculated (Ranstam et al. 2000).

In the early days of RSA, the accuracy of rigid body kinematics, derived from phantom studies, was quoted at 0.01 to 0.25 mm for translations and 0.03 to 0.6 degrees for rotations (Kärrholm 1989). Subsequently the accuracy of the RSA method has been improved even further, for example, with the introduction of digital measurements of the radiographs (Börlin et al. 2002). Indeed, at least in an idealized *in vitro* simulation-study, sub-10 µm accuracy has been reported (Madanat et al. 2007). Moving away from ideal laboratory conditions additional error sources are introduced and the measurement accuracy is reduced. As a demonstration of this eventuality, in a human cadaveric study of a hip stem implant, accuracy of 0.047 to 0.12 mm has been reported (Önsten et al. 2001).

Table 1. The precision or zero-motion confidence intervals reported in recently published RSA studies of either femoral stem or acetabular cup migration.

				Reported zero-motion 95% CI	
Year	**Authors**	**RSA method**	**Implant**	**Translation**	**Rotation**
2020	Tabori-Jensen et al.	Model-based	Acetabular cup	0.18 to 0.39 mm	1.25 to 1.8 degrees
2020	Kruijntjens et al.	Model-based	Femoral stem	0.24 to 0.64 mm	0.18 to 1.49 degrees
2020	Jørgensen et al.	Model-based	Acetabular cup	0.21 to 0.61 mm	0.36 to 1.56 degrees
2020	Floerkemeier et al.	Model-based	Femoral stem	0.06 to 0.23 mm	0.21 to 3.12 degrees
2020	Howie et al.	EGS	Acetabular cup	0.01 to 0.08 mm	0.11 to 0.43 degrees
2020	Reiner et al.	Model-based	Femoral stem	0.17 to 0.62 mm	0.42 to 1.29 degrees
2020	Thoen et al.	Marker-based	Acetabular cup	0.08 to 0.12 mm	0.10 to 0.12 degrees
2020	van der Voort et al.	Marker-based	Femoral stem	0.17 to 0.54 mm	0.31 to 1.13 degrees
2020	Bergvinsson et al.	Marker-based	Acetabular cup	0.12 to 0.39 mm	0.55 to 1.38 degrees
2020	Dyreborg et al.	Model-based	Femoral stem	0.15 to 0.54 mm	0.25 to 2.32 degrees

Already in the 1990s the clinical precision of marker-based RSA in THA studies was quoted at between 0.15 to 0.6 mm for translation and 0.3 to 3 degrees for rotation (99% CI of containing a double measurement) (Kärrholm et al. 1997). The reported 95% confidence intervals in recent literature are remarkably in-line with this figure (**Table 1**). However, various error sources still exist that can degrade the precision of the method and few RSA setups are technically equivalent. Therefore, routine double examinations to confirm the clinical precision of each RSA setup have been recommended (Valstar et al. 2005).

Mean Error of Rigid-Body Fitting and Condition Number

When studying the motion of rigid-bodies additional potential sources of error are added into the system when compared to analyzing only individual RSA markers (Selvik 1989). Non-conformity to the rigid body assumption and poor spatial distribution of the RSA markers can degrade the precision of RSA measured rigid-body kinematics (Ryd 1986, Söderkvist and Wedin 1993). It is therefore necessary for the researcher to have methods to evaluate the quality of rigid body data and the proper scatter of RSA markers.

Mean error of rigid body fitting (ME) has been established as the standard method for assessing stability of RSA markers (Valstar et al. 2005). Originally in his thesis Selvik used ME as the cost function for optimizing the rigid body transformation itself (see Box 2) (Selvik 1989). As described by Selvik, ME is calculated as the root mean square distance between observed marker positions and calculated marker positions derived from an earlier RSA examination by eliminating the estimated rigid body transformation between the examinations. Understandably, errors in the estimated rigid body transformation itself will lead to increased ME. Besides errors in rigid body transformation, the other major source of ME is the elasticity of the rigid body, i.e. non-conformity to the rigid body assumption. Assuming ME is indeed minimized by the relevant optimization algorithm, only the latter error source will significantly contribute to ME. In other words, the lower bound for the resulting ME is dictated by marker stability between any two RSA examinations.

As instability of markers in a rigid body will lead to non-conformity to the rigid body assumption, it may compromise the calculation of rigid body transformations. Indeed, Leif Ryd (1986) showed that increasing ME as a result of induced errors in individual marker positions was associated with increasing errors in RSA measured migration. Consequently, an upper limit of 0.35mm for ME has been recommended for assessing sufficient marker stability in the RSA guidelines (Valstar et al. 2005).

ME alone is not sufficient to ensure a good configuration of RSA markers. Beyond marker stability, the proper scatter of markers is important especially when

rotations are calculated. This is evident even in the simplified case of calculating the angle of a line between two RSA markers. The angle between the markers is relatively well defined when the distance between them is considerably larger than the uncertainty in the markers' measured position. On the other hand, if the markers are spaced close to together, the measurement error in their absolute position may become a considerable determinant of the calculated angle. Thus, even rigid bodies with stable markers and corresponding low ME may have poor precision of RSA measured rotation. This is true for all rigid bodies where any number of markers are configured close to a line (collinearity) or close to a single point.

Consequently, to complement ME, c*ondition number* (CN) has been recommended for the evaluation of proper marker scatter (Söderkvist and Wedin 1993, Ryd et al. 2000, Valstar et al. 2005). CN can be calculated by first obtaining a least square fit line through the markers in a rigid body. CN is then the inverse of root square sum of distances between the line and each marker in the rigid body. Therefore increasing CN corresponds to increasing collinearity of RSA markers and a compromised precision of determining rotation around said line (Ryd et al. 2000). Together, ME and CN can be used then to estimate the level of measurement noise of both translation and rotation in an RSA measurement. An upper limit for acceptability of 150 has been suggested for the CN (Valstar et al. 2005). Especially in studies involving small joints larger CNs may be considered acceptable: with smaller joints the physical limits of the bone structures often do not allow for ideal scatter of RSA markers and concessions in the precision of the rotational measurements have to be made (Valstar et al. 2005).

Statistical Power

Minimizing the sample size while maintaining adequate statistical power is a critical consideration for all well-designed and ethically sound clinical research (Vollmer and Howard 2010). A smaller sample size translates to a smaller number of study participants subjected to the potential of harm due to any clinical intervention and, usually, to more efficient use of limited research resources (financial or otherwise). Meanwhile, underpowered research may be viewed as exposing the study participants to unnecessary harm while wasting the associated resources without a tenable promise of any gained knowledge (Maxwell and Kelley 2011). Underpowered research may even negatively impact a whole research field as seemingly contradicting results, representing merely sampling error, are published (Maxwell and Kelley 2011). Thus, well-designed clinical research must strike a careful balance between as small as possible but sufficiently large sample size.

For RSA studies, astonishingly small sample sizes of 15 to 25 patients per group have been suggested for making reliable long-term predictions of implant revision

rates due to aseptic loosening (Valstar et al. 2005). We can confirm this assessment by utilizing data from systematic reviews on RSA. In the case of the cemented femoral stem, a two-year stem subsidence of more than 0.15mm is considered an indication of possible revision rates of more than 5% at ten years (van der Voort et al. 2015). In the case of, for example, the cemented Lubinus SP II femoral stem the pooled standard deviation for subsidence at two years was 0.26 mm (van der Voort et al. 2015). If we then wanted to detect group-level subsidence of greater than 0.15 mm with a power of 80% and at an α-level of 5%, a minimum group size of 19 would be required. Similarly, for a two-sided between groups analysis a per-group size of 48 would be required if a value as low as 0.15 mm was to be assumed as a clinically relevant between-groups difference.

For the acetabular cup a two-year subsidence suggestive of greater than 5% revision rates was identified at 0.2mm (Pijls et al. 2012a). In the same systematic review, the pooled standard deviation was 0.24mm for the cementless porous-coated Harris-Galante I acetabular cup. With identical power and α-levels a minimum group size of as low as 18 is calculated. For a two-sided between groups analysis a sample size of 23 patients per group is calculated for detecting a difference in subsidence of 0.2mm.

An alternative power calculation for RSA studies has been provided by Derbyshire et al. (2009). In their work it was argued that the previously recommended small sample size of 15-25 patients is a gross underestimation. They argued that such small sample sizes will result in excessive confidence intervals for the predicted population level revision rates. This claim was based on the argument that only a few percent of implants are eventually revised by 10 years and that it is therefore likely that such failure cases would likely be missed by a small sample of only 15-25 patients. Therefore, Derbyshire et al. (2009) recommended a much larger minimum sample size of 46 patients per group for RSA studies. The logical fallacy of this argument is, however, that in their calculations Derbyshire et al. (2009) essentially assume that only implants revised by 10 years exhibit excessive early migration. Yet, it has been shown that, at least in the case of cemented acetabular cups, only a fraction of the implants that show increased early migration undergo revision by 10 years (Aspenberg et al. 2008). Thus, the sub-population of implants that show increased early RSA-measured migration is likely much larger than the number of patients that will undergo revision by 10 years. This seems to support the argument that the smaller sample size of 15-25 patients in a typical RSA study is justified and can be used to make reliable estimates of the population level revision rates even if no early revision cases are included in the RSA study.

2.2.5 Causalities of RSA-Measured Micromotion

Worldwide, aseptic loosening is the leading indication for revision surgery after primary THA (Ulrich et al. 2008, Sadoghi et al. 2013, Fernández-Fernández et al. 2020). The etiology of aseptic implant loosening is likely multifactorial. Suspected or known risk factors for aseptic loosening include wear particles (polyethylene, ceramic, or metal), stress shielding, adverse reactions to bone cement, patient related factors, high fluid pressures at the bone-implant interface, and inducible micromotion. Many of these risk factors are likely also causally interlinked (Sundfeldt et al. 2006).

The association between two-year RSA-measured micromotion and 10-year revision rates due to aseptic loosening have been confirmed in systematic reviews for various types of orthopaedic implants (Pijls et al. 2012a, 2012b, van der Voort et al. 2015). A perhaps classical view is that the early migration is a sign of the one and same process – aseptic loosening – only at an earlier stage (Pijls et al. 2012b, Mjöberg 2020). Still, from the viewpoint of causalities it is important to make the distinction that a static RSA measurement of *sustained* micromotion – as opposed to cyclic or inducible micromotion – is not a direct measure of loosening (**Figure 4**). It is more likely a surrogate measure ("confusion factor") for the various processes that ultimately lead to the clinical manifestation of aseptic loosening (Kundi 2006, Sundfeldt et al. 2006, Dodge 2008).

The theory that sustained micromotion is a common denominator for the various causal pathways to aseptic loosening underlies much of the RSA literature. Considering the accumulated evidence of an association this concept seems well supported (Pijls et al. 2012a, 2012b, van der Voort et al. 2015). However, it is conceivable that not all processes that lead to aseptic loosening manifest as increased early migration in RSA studies. For example, an inferior implant design may not migrate and become loose until after the typical 2-year follow-up period due to delayed mechanical or biological processes. Additionally, implants may ultimately fail for reasons wholly unrelated to aseptic loosening (Sköldenberg et al. 2014, van der Voort et al. 2015). Conversely, as has been observed in the case of individual implant types, not all micromotion necessarily leads to aseptic loosening (Nieuwenhuijse et al. 2012). With these considerations in mind, it is paramount to confirm RSA-predicted implant survival with the help of other research methods such as implant registries and clinical prospective cohort studies (Derbyshire et al. 2009, Frazer and Tanzer 2020).

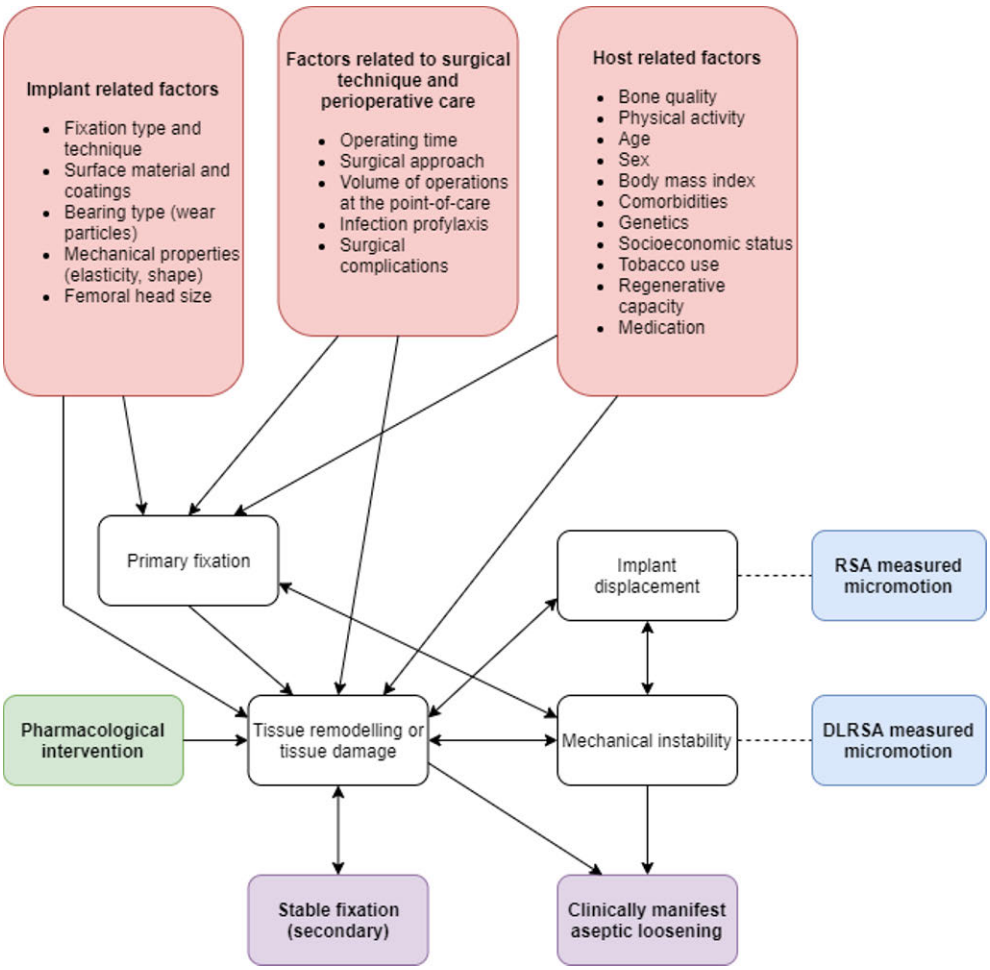


Figure 4. A causal diagram of risk factors and protective factors for aseptic loosening of THA (red) in relation to RSA-measured micromotion (Sundfeldt et al. 2006, Prieto-Alhambra et al. 2011, Prokopetz et al. 2012, Mellon et al. 2013, Shi et al. 2018).

Cyclic or inducible micromotion may play a more central role in the causal pathway to aseptic loosening and may even be viewed as having a direct causal relationship with aseptic loosening (**Figure 4**). Cyclic micromotion at the bone-implant interface can be detrimental for the fixation of either cemented or cementless implants. With cementless implants the issue is more clear-cut: The main principle of cementless THA, for example, is to achieve osseointegration with the host bone (Spotorno et al. 2012). To enable osseointegration, a sufficiently rigid initial fixation is paramount. Cyclic micromotion of as low as 40 µm to 150 µm at the bone-implant interface may inhibit bone growth onto and into the implant (Ryd et al. 1993, Goodman 1994, Jasty et al. 1997). It then follows that implants displaying excessive inducible micromotion

cannot be osseointegrated whereas a sufficiently rigid mechanical connection beyond the immediate postoperative period implies stable osseointegration (Ryd et al. 1990, Laende et al. 2019b).

The significance of inducible micromotion for cemented implants is a more complex issue. In addition to direct bone contact, long-term fixation of a cemented implant may be characterized by fibrous tissue and fibrocartilage tissue in-between the bone and the implant (Giori et al. 1995). It stands to reason, then, that somewhat greater inducible micromotion may be observed for a cemented implant (Mann et al. 2012). Still, excessive inducible micromotion may prevent or compromise mechanical interlocking of bone and cement undermining the fixation stability. For both cemented and cementless implants, cyclic micromotion may also allow wear particles to travel between the implant-bone interface and even generate high hydrostatic pressures ultimately leading to adverse tissue reactions (Sundfeldt et al. 2006).

These conceptual and causal differences of fixation between cemented and cementless implants help in understanding why sustained micromotion, inducible micromotion, and their timing may have differing impacts for these fixation types. While osseointegration of cementless implants requires stable fixation, some cemented THA components have been designed to subside and show good clinical long-term outcome, regardless (Nieuwenhuijse et al. 2012). It has been noted that, besides different fixation types, the migration pattern may also change from one implant design to another or even according to surgery- and patient-related factors (de Vries et al. 2014). Indeed, various implant-related and surgical factors have been linked to both implant primary fixation and host tissue response (**Figure 4**).

Compared to the study of orthopaedic implants, relatively few RSA studies of fracture healing have been performed to date. Consequently, the causalities of RSA measured micromotion are perhaps less well defined for the healing of bone fractures. Some parallels between the biology and biomechanics of fracture healing and fixation of cementless implants can be drawn – especially concerning primary fracture healing which is the aim of stable osteosynthesis (Marsh and Li 1999, Marsell and Einhorn 2011). It follows that the key causalities between RSA-measured micromotion and fracture healing likely resemble those concerning cementless implant fixation (**Figure 4**). Indeed, a link between RSA-measured interfragmentary micromotion and an inferior clinical outcome has been demonstrated at least for some fracture types (Madanat et al. 2006).

Differences in implant fixation and bone fracture healing become more pronounced when considering the process of secondary fracture healing seen with non-rigid fixation (Marsh and Li 1999, Marsell and Einhorn 2011). For secondary fracture union, cyclical axial micromotion at the fracture site is a necessary phenomenon to promote the formation of an external callus whereafter a slow and

progressive stabilization and strengthening at the fracture-site occurs (Marsh and Li 1999). Thus, secondary fracture healing is not a dichotomous event but is rather characterized by a carefully organized progression of biological and biomechanical processes (Marsell and Einhorn 2011, Bizzoca et al. 2020). This translates to gradually increasing stiffness and strength of the fracture with no clear-cut distinction between union and non-union (Richardson et al. 1994, Chehade et al. 1997, Claes and Cunningham 2009, Morshed 2014). Meanwhile, delineating the stage of fracture healing has proven an exceedingly difficult task *in vivo* and an issue that hampers the related clinical research (Morshed 2014).

Considering these concepts of bone healing, DLRSA seems uniquely positioned as a research tool for the study of fracture union. DLRSA has already been demonstrated for the direct measurement of fracture-site stiffness which in turn has been shown to correlate with the increasing strength of the healing fracture (Chehade et al. 2009, Claes and Cunningham 2009). On the other hand, in situations where primary union is desired, DLRSA may be a useful tool for confirming sufficient early stability after surgical reduction and fixation. However, the clinical data gathered thus far is still very limited and these hypothetical use-cases of DLRSA and the causalities of inducible micromotion still lack validation.

2.2.6 Competitive Methods

Conventional Radiography

In the clinical setting, due to its availability and low cost, plain radiography is still considered a cornerstone in the postoperative assessment of orthopaedic implants (Vanrusselt et al. 2015, Cyteval 2016). However, conventional radiography can often provide only very crude estimates of implant migration: in a study comparing RSA and plain radiography, the accuracy of plain radiography (95% CI of difference) varied between 3.9 to 12.3mm for the measurement of hip stem migration depending on the used landmarks (Malchau et al. 1995). Similarly, for the acetabular cup the accuracy of plain radiography was only 4.4 to 6.5mm (Malchau et al. 1995). As an indirect result of the poor measurement accuracy, the significance of secondary findings, like radiolucencies, is often emphasized when analyzing plain radiographs (Vanrusselt et al. 2015, Cyteval 2016).

The use of conventional radiography for the follow-up of fracture healing is commonplace in the clinical research. This is despite the fact that, at least for metaphyseal fractures, the findings provided by conventional radiography seem to correlate poorly with the state of fracture healing (Morshed et al. 2008, Chehade et al. 2009, Claes and Cunningham 2009). Additionally, conventional radiography has

been shown to perform poorly in assessing the biomechanical properties of the bone (McClelland et al. 2007).

Even fracture displacement can be assessed only very crudely with plain radiographs: for femoral neck fractures the accuracy (95% CI of difference compared RSA) of measuring femoral head displacement has been reported at only 5.8 to 9.6 mm depending on the direction (Ragnarsson et al. 1992). As a result, secondary findings are emphasized over fracture-site displacement as outcome measures in plain radiographs for clinical research (Morshed et al. 2008).

Computer-Aided Analysis of Radiographs

Although RSA is considered a relatively safe method, evidence for clinical use of RSA outside dedicated research is still lacking (Shah et al. 2018). Consequently, for the routine clinical practice, various computer-aided methods aimed at measuring implant migration from conventional radiographs have been developed. Among others, these methods include Einzel Bild Röntgen Analyse (EBRA), Düsseldorf Migration Analysis, Ulm Migration Analysis and Manchester X-ray Image Analysis for the hip implant (Schütz et al. 2005).

The computer-aided methods primarily rely on estimating and correcting for changes in the orientation of the implant in serial radiographs. This accommodates more reliable measurements of migration compared to manual measurements from conventional radiographs. However, these methods are limited in the degrees of freedom for which the migration can readily be measured (Schütz et al. 2005). Also, even one of the most accurate of these computer-aided methods, EBRA, has shown measurement accuracy of only 0.8 to 1.6mm for THA components in relation to bone (Schütz et al. 2005, Malak et al. 2016).

While the accuracy quoted for EBRA may numerically seem quite close to RSA, the implications for clinical research of THA components, for example, are quite stark (**Box 3**). Still, although the accuracy EBRA does not compare to that of RSA, considering the relative simplicity of the method, the improvement in accuracy over manual measurements may be useful in the daily clinical practice and some clinical research where RSA cannot be applied.

BOX 3. The effect of diminishing accuracy on statistical power

Using the Harris-Galante I acetabular cups (also used in the power calculation for RSA previously) as an example, we can estimate the sample size needed in an EBRA study as compared to RSA. We will aim to detect a sample mean migration exceeding 0.2mm at two years. The key parameter for a power calculation is the standard deviation of the EBRA measurements. However, this value is not readily available for EBRA of the Harris-Galante I acetabular cups at two years. We can still estimate the lower bound for the standard deviation in the following manner:

We will assume insignificant systematic error for the benefit of EBRA. Then, the remaining variance of EBRA measurements consists of the sample variance of true migration and random error of the measurement technique. The variance of random error in an EBRA measurement of acetabular cup subsidence has been estimated at 0.12 mm^2 (Krismer et al. 1995). Assuming RSA has no significant bias, we can approximate the sample variance of true migration from previous RSA publications. Given the RSA measured standard deviation at two years for the Harris-Galante I acetabular cups and the random error of the corresponding RSA measurements (through double measurements of zero-motion) we can calculate the 2-year sample variance of true migration at roughly 0.05 mm^2 (Önsten and Carlsson 1994, Önsten et al. 1994, Pijls et al. 2012a).

Since variances of the error terms can be added (by assuming no correlation between the population and random error variances) we arrive at an estimated best-case variance of 0.05 mm^2 + 0.12 mm^2 = 0.17 mm^2 for an EBRA measurement of a Harris-Galante I acetabular cup at 2 years post operatively. The corresponding sample standard deviation would be 0.41 mm. This figure is remarkably in-line with the standard deviation of 0.53 mm that has been observed in a clinical EBRA study of Harris-Galante I acetabular cups at 65 months mean follow-up. Using an alfa-level of 0.05 and statistical power of 80%, the corresponding minimum group size would be 66 (sample SD of 0.41 mm) for detecting a between-group difference exceeding 0.2mm – an almost three-fold increase compared to RSA, as calculated previously. In a real-world study, considerations such as possible bias introduced by EBRA would likely inflate the sample-size even further.

Computed Tomography

With CT the principles of rigid body kinematics can be applied. Recent advances in the related methodology have resulted in measurement accuracy and precision comparable to RSA with or even completely without tantalum markers (Olivecrona et al. 2016, Sandberg et al. 2020). The method relies on volumetric feature matching

of bone, markers, and the implant in serial CT acquisitions akin to what has been accomplished with RSA using CT generated surface models (Olivecrona et al. 2016).

In the marker-free CT-based techniques the measurement accuracy of implant migration has been reported at 0.28 to 0.7 mm for the acetabular cup in phantom studies (Olivecrona et al. 2003, Scheerlinck et al. 2016). Similarly, in a porcine cadaver study of marker-free CT analysis, the relative motion of the femoral head and acetabular cup, i.e. the liner wear, could be determined with an accuracy of 0.5 mm (Sandgren et al. 2016). Using markers may improve measurement accuracy even further. Brodén et al. (2016) found the measurement accuracy of a marked acetabular cup phantom to be 0.07 to 0.32 mm for translation and 0.21 to 0.82 degrees for rotation using CT. In the same study no significant differences in precision could be detected between a CT analysis and a simultaneous RSA analysis. Olivecrona et al. (2016) reported even better accuracy for CT migration analysis – albeit in a highly idealized study setting – at under 0.16 mm on all axes on an acetabular cup phantom with tantalum markers.

The clinical precision for a CT-based method, termed CT-based implant motion analysis (CTMA), was only recently reported at 0.1 to 0.3 mm and 0.1 to 0.4 degrees (95% CI of zero-motion) for the femoral stem (Sandberg et al. 2020). For the acetabular cup, CTMA provided clinical precision of 0.07 to 0.31 mm and 0.2 to 0.39 degrees for rotation. Using markers did not improve the measurement precision considerably in the clinical setting (Brodén et al. 2020). These precision values are well within those reported in RSA studies (**Table 1**). Interestingly, metal artefacts – inherent to CT imaging of radiodense orthopaedic implants – were an apparent non-issue in these studies even when not utilizing the RSA-markers (Brodén et al. 2020, Sandberg et al. 2020).

The main issue hampering the deployment of CT-based migration analysis has been the high radiation dose (Sandgren et al. 2016). In the micromotion study of the hip stem, Scheerlinck et al. (2016) reported an average effective dose per CT acquisition of 5.5 mSv. During the same year Olivecrona et al. (2016) reported effective radiation doses at between 2.4 to 6.6 mSv per CT acquisition for their CT-based implant motion analysis. Fortunately, improving equipment and methodology have seen the estimated effective doses decreasing in more recent studies. For the CTMA-method, Brodén et al. (2020) reported a mean calculated dose of 0.2 to 2.3 mSv per CT acquisition depending on the study center. Still, these numbers are in stark contrast with RSA and even conventional radiography of the hip. The corresponding radiation doses have been calculated at 0.15 mSv and 0.28 mSv for RSA and conventional radiography, respectively, in the study of the hip joint (Valstar 2001). Quite recently, in a phantom study of the hip, a study protocol for model-based RSA was described with as low effective dose as 0.043 mSv per radiostereograph acquisition (Blom et al. 2020).

The fairly universal availability of CT over specialized RSA setups may serve to promote the CT-based micromotion analysis methods (Otten et al. 2017). Also, the CT-based analysis methods don't suffer from the marker occlusion – a recognized issue in the RSA studies of the acetabular cup, for example (Brodén et al. 2016). For specialized research, the ability to use the same set of CT acquisitions for migration analysis and diagnostic purposes (of, for example, periprosthetic osteolysis) could also be beneficial (Sandgren et al. 2016). However, the RSA methodology is also evolving and thereby limiting the use-cases where the comparatively high radiation dose of CT can be overlooked. For example, the possibility of marker-free RSA, covered previously, effectively does away with the issue of marker occlusion.

2.3 Clinical Applications of RSA in Cementless THA

2.3.1 The Coevolution of Cemented and Cementless THA

The earliest documented attempts at hip arthroplasty date back to the late 19th century (Learmonth et al. 2007). Still, it wasn't until the 1960s that total hip arthroplasty saw its first great success with polymethylacrylate cement fixation introduced by Sir John Charnley (Charnley 1961). The 10-year and even longer-term results of the procedure, as introduced by Charnley, were by and large considered excellent (Callaghan et al. 2009). Besides the choice of the cementing medium, another key factor for the success of Charnley's total hip arthroplasty was the use of a small diameter metal femoral head in conjunction with a polyethylene acetabular cup. The small diameter femoral head minimized the frictional torque and wear at the articulation. Further, the low friction properties and relative biocompatibility of polyethylene lent themselves for the long-term success of the Charnley arthroplasty construct (Charnley 1961, Learmonth et al. 2007).

However, even Charnley himself recognized that the longevity of THA was definite and later cautioned against the procedure without a compelling clinical indication (Charnley 1961, 1970). Early on, concerns were raised over the possibility of adverse biological reactions to the bone cement – especially among the younger patient demographics (Willert et al. 1974, Jones and Hungerford 1987). The perceived risks of cement use, whether real or not, drove the development of cementless techniques. With cementless THA the fixation of the implant to the host bone was sought through *biological fixation* – i.e. direct osseointegration of the implant with host bone (Lord et al. 1979, Spotorno et al. 2012). Although experiments with cementless implants had been conducted previously (McKee and Watson-Farrar 1966), cementless THA started to gain traction only in the late 70s

and saw increasing clinical application from the early 80s onwards (Spotorno et al. 2012).

Early attempts at cementless THA during the 1960s and 1970s were plagued by high revision rates. At the time, cementless implants often had a smooth surface and employed macroscopic augments aimed at providing stability (Yamada et al. 2009). These macroscopic augments included, for example, collars and holes for the femoral stem and large screws or even fully threaded external surfaces for the acetabular cup (Ring 1968, Lord et al. 1979). However, instead of the macroscopic, the key breakthrough for the cementless THA came at the microscopic level. During the late 1970s and early 1980s it was shown that the bone ingrowth and, subsequently, the strength of the bone-implant interface could be improved with a porous implant surface finish (Robertson et al. 1976, Bobyn et al. 1980, Spector 1987). With porous-coated implants very high success rates of osseointegration could be achieved for both the femoral stem and the acetabular cup (Engh et al. 1990). In a sense, the introduction of porous coatings marked the birth of modern cementless THA with most contemporary implants employing some type of porous coating (Yamada et al. 2009).

Meanwhile, the expected outcome of cemented THA has been improving, as well, owing largely to the refinement of cementing techniques (Niculescu et al. 2016). Both cemented and cementless techniques have benefitted from the continued development of improved polyethylene bearing surfaces with reduced rates of wear (Shi et al. 2019, Wyatt et al. 2019). At present, both fixation methods and their per-component combinations can provide excellent long-term clinical results with neither technique having a categorical advantage over the other (Moskal et al. 2016, Konan et al. 2019, Praet and Mulier 2019). Still, for individual patients and patient demographics the choice of the fixation method is not irrelevant. Therefore, a patient-by-patient consideration of the optimal fixation method has been recommended (Konan et al. 2019). Such individualized decision-making requires a close understanding of how the various patient-related factors affect the risk of revision for either fixation type.

The Decreasing Risk of Revision and the Role of RSA

As the risk of revision – particularly due to aseptic loosening – after THA has been decreasing (Fernández-Fernández et al. 2020), it might at first be surmised that the value of RSA as a research tool for THA would have been equally diminished. However, somewhat unintuitively, the principal advantage of RSA as a research tool in THA does not lie in explicitly minimizing the risk of aseptic loosening of a particular implant component. Instead, RSA has been envisioned a key role in the

future development of THA as part of a phased clinical introduction scheme for new implants (Nelissen et al. 2011, Pijls and Nelissen 2016).

The ability of RSA to predict the long-term revision risk of most THA components from relatively small patient cohorts with only a two-year follow-up make for a highly ideal screening tool. In phased clinical introduction of new implants, RSA screening studies could be used as a gate-keeper step to prevent implants with unexpectedly high rates of revision from entering more widespread clinical use (Nelissen et al. 2011, Pijls and Nelissen 2016). Indeed, as discussed previously, improving patient safety is one of the main arguments for the phased introduction of THA components (Pijls and Nelissen 2016). However, the advantages of RSA as a screening tool go beyond this gate-keeper role. RSA also has the potential to accelerate the development of new THA implants and save valuable research resources. With RSA, a poor performing design can be identified and abandoned forthwith without the need to wait up to decades for the results of conventional clinical follow-up studies involving up to hundreds of patients (Michelson and Riley 1989, Nelissen et al. 2011). Thus, although the relative significance of aseptic loosening as a cause of revision may have decreased over the years, the importance of RSA as an enabler of safe innovation in the field of THA has only been emphasized.

2.3.2 The Acetabular Cup

Although the ground work for RSA of the acetabular cup was laid already in 1980, it wasn't until 1990 that the first RSA study on a cementless acetabular cup was published (Baldursson et al. 1980, Snorrason and Kärrholm 1990). Thereafter a large number of original RSA studies have been published on the cementless acetabular cup (**Table 2**). Even more importantly, with the aid of data derived from arthroplasty registers, the relationship between 2-year RSA-measured proximal migration and 10-year risk of revision for aseptic loosening was established in a meta-analysis only underlining the significance of earlier pioneering RSA studies of the cementless cup (Pijls et al. 2012a).

An issue relevant to RSA studies of the acetabular cup (**Table 2**) is the significant number of dropouts. In studies with a planned two-year follow-up the mean dropout rate was 15% (range 2 to 46 %) at the last follow-up. A recurring reason for losing a cup to follow-up was technical difficulties associated with occluded markers either on the acetabular cup or the surrounding bone.

Table 2. Clinical RSA research published on migration of cementless acetabular cups.

Year	Authors	RSA Study setting	n (male/female)	RSA follow-up
2020	Bergvinsson et al.	Liner wear (and cup migration) according to femoral head type (RCT)	50 (34/16)	5 years
2020	Howie et al.	Comparison of titanium cups with and tantalum cups without screws (RCT)	66 (35/31)	2 years
2020	Tabori-Jensen et al.	Comparison of cementless vs cemented cups in the elderly	30 (13/17)	2 years
2020	Laende et al.	Cohort study	29 (13/16)	3 years
2020	Jørgensen et al.	Porous vs porous hydroxyapatite coated cup (RCT)	53 (25/28)	2 years
2018	Jacobsen et al.	Cohort study (comparing EGS-RSA and RSA)	50 (NA/NA)	10 years
2017	Nilsson et al.	Cohort study	20 (8/12)	5 years
2017	Mohaddes et al.	Cemented vs uncemented revision (RCT)	20 (11/9)	17 years
2017	Shareghi et al.	Retrospective comparison of MBRSA with RSA on cementless cups	80 (NA/NA)	2 years
2016	Minten et al.	Effect of screw fixation on migration and liner wear (RCT)	37 (15/22)	6.5 years
2016	Otten et al.	Long-term effect of fixation augments (screws, pegs etc.) (RCT)	68 (NA/NA)	14 years
2015	Ayers et al.	Comparison of two cementless cups and liner materials (RCT)	46 (19/27)	5 years
2014	Lazarinis et al.	Cohort study	30 (12/18)	2 years
2014	Saari et al.	Effect of risedronate on cup stability after revision (RCT)	27 (NA/NA)	3 years
2013	Naudie et al.	Comparison of titanium cup surface finish (RCT)	62 (19/43)	2 years
2013	Munzinger et al.	Effect of hydroxyapatite-coating on migration in females (RCT)	44 (0/44)	2 years
2012	Pakvis et al.	Comparison with and without screw fixation (RCT)	37 (15/22)	2 years
2012	Wolf et al.	Effect of early weight-bearing on migration (RCT)	30 (14/16)	5 years
2011	Baad-Hansen et al.	Comparison of two cementless cups (RCT)	60 (34/26)	2 years
2010	Wolf et al.	Study on the timing of the first RSA acquisition (RCT)	24 (14/10)	7 days
2007	Thien et al.	Effect of early weight-bearing on migration (RCT)	43 (20/23)	1 year
2006	Zhou et al.	Effect of liner on migration (RCT)	61 (32/29)	2 years
2006	Carlsson et al.	Comparison of a novel cementless THA with established designs (RCT)	40 (24/16)	2 years
2006	Carlsson et al.	Comparison of a novel cementless THA with established designs (RCT)	53 (NA/NA)	3 years

Year	Authors	RSA Study setting	n (male/female)	RSA follow-up
2004	Röhrl et al.	Effect of fixation augments (screws, pegs etc.) on migration (RCT)	87 (44/43)	5 years
2004	Digas et al.	Cementless vs cemented cups migration and liner wear (RCT)	37 (8/29)	2 years
1998	Önsten et al.	Cementless vs cemented cup liner wear and migration (RCT)	51 (NA/NA)	7 years
1997	Nivbrant et al.	Cohort study	43 (17/26)	2 years
1996	Thanner et al.	Comparison of screw material (RCT)	43 (17/26)	2 years
1996	Nivbrant et al.	Cohort study	60 (36/24)	2 years
1995	Önsten et al.	Cohort study (migration correlated with local bone morphology)	19 (11/8)	2 years
1992	Kärrholm et al.	Cohort study	22 (16/6)	2 years
1990	Snorrason et al.	Cohort study	20 (8/12)	2 years

NA; not available.

As discussed previously, cyclic micromotion at the bone-implant interface exceeding as little as 40 μm may prevent the successful osseointegration of a cementless implant. It is understandable then that the importance of close implant-bone opposition and stability of initial fixation have been emphasized for the success of cementless THA (Illgen and Rubash 2002, Spotorno et al. 2012). Consequently, factors aimed at maximizing initial stability, such as adjunct fixation methods and coatings that may promote osteoconduction, represent a popular topic in the RSA research of the cementless cup (**Table 2**). The effect of host bone quality on initial implant stability has seen only meager interest among the RSA studies of the cementless cup (Önsten et al. 1995).

Despite the fact that good bone quality was considered a prerequisite by the early pioneers of cementless THA (Spotorno et al. 2012), the original age indication has only expanded to include older patient populations (Dutton and Rubash 2008). This development has been called “the uncemented paradox” and it has been emphasized that cemented THA should be considered the primary option for patients over 75 years of age (Troelsen et al. 2013). At the same time the prevalence of compromised bone quality may have been much higher among the patients considered for THA than previously thought – even among patients under 75 years of age (Glowacki et al. 2003, Mäkinen et al. 2007). Indeed, at least for the uncemented femoral stem, low BMD is associated with increased RSA-measured migration (Sköldenberg et al. 2011, Aro et al. 2012). Yet, there is an apparent paucity of research on the effect of systemic skeletal status on the cementless acetabular cup.

2.3.3 The Femoral Component – Implant Surface MBRSA

RSA has time and again been criticized for being time consuming and expensive. As discussed previously, one major issue facing RSA researchers even before starting a clinical trial is obtaining implants that have been marked with RSA beads (Valstar et al. 2001, Kaptein et al. 2003). Therefore, it is obvious that model-based RSA has the potential to considerably accelerate and enable the study of new implant designs. The implant surface model-based methodology for the femoral stem has recently matured from methodological studies to clinical research as is evident from increasing number of clinical studies using MBRSA (**Table 3**). Various types of femoral stems have been successfully used in these studies underlining the flexibility of the method. These studies have, also, not reported on any major issues with the model-based analysis.

Table 3. MBRSA research focusing on the femoral stem. Only surface-model –based studies included.

Year	Authors	RSA Study setting	n (male/female)	Follow-up
2020	Perelgut et al.	Comparison of collared and collarless stems (RCT)	58 (NA/NA)	1 year
2020	Dyreborg et al.	Comparison of two stems (RCT)	62 (34/28)	2 years
2020	Reiner et al.	Comparison of two stems (RCT)	44 (15/29)	2 years
2020	Floerkemeier et al.	Effect of 3-month migration on RSA/clinical outcome	60 (26/34)	5 years
2020	Nieuwenhuijse et al.	Comparison of two stems (RCT)	51 (16/35)	2 years
2020	Christiansen et al.	Cohort study	50 (45/5)	2 years
2019	Richardson et al.	Cohort study	25 (4/21)	2 years
2018	Hoornenborg et al.	Effect of a coating on migration (RCT)	51 (20/31)	2 years
2018	Schwarze et al.	Effect of surgical approach (RCT)	60 (26/34)	2 years
2017	Floerkemeier et al.	Factors influencing short stem migration	78 (NA/NA)	2 years
2016	Budde et al.	Cohort study	18 (4/14)	2 years
2016	Acklin et al.	Cohort study	34 (NA/NA)	2 years
2014	Hjorth et al.	Metal ions vs stem/cup migration	49 (NA/NA)	5 years

NA; not available.

2.4 Clinical Applications of DLRSA in Fracture Healing Studies

The application of RSA for the study of fractures has arguably been overshadowed by the enthusiasm surrounding the study of early implant micromotion. Only

relatively few RSA studies on micromotion of traumatic fractures have been published (**Table 4**). At the same time, there seems to be a lack of established methods for defining fracture union *in vivo* (Bhandari et al. 2002, 2013, Morshed et al. 2008, Kooistra et al. 2010, Morshed 2014). Indeed, the lack of proper methods for establishing fracture-union has been recognized as a major obstacle for the related orthopaedic research (Corrales et al. 2008, Augat et al. 2014, Morshed 2014, Bizzoca et al. 2020).

Not until quite recently, DLRSA has been shown to provide unique possibilities for evaluating the biomechanics of fracture-healing *in vivo.* Previously, inducible RSA-measured migration has been utilized to characterize the mechanical properties of the fracture-site by only few authors (**Table 4**). Since only very few DLRSA studies on bone fractures have been published, the methodology has not yet been thoroughly established. For many fracture-sites, like the hip, no previous DLRSA studies have been performed. As discussed previously, standardizing and controlling for the force that is exerted at the fracture-site during DLRSA is crucial. Thus, methods for generating the differential loading at the hip are yet to be developed and validated.

Similarly, clinically relevant cut-offs for stiffness indicative of union or non-union are by and large yet to be defined. Complicating the matters further, it has been noted that, when external forces are applied, the elasticity of the implant and/or the bone tissue will result in micromotion observable with RSA (Ryd et al. 1993). This motion may be detected with RSA even if no micromotion were present at the bone-implant interface (Wilson et al. 2010). In the case of tibial total knee arthroplasty this micromotion has been estimated at up to 0.3mm of maximum total point motion (MTPM) (Little et al. 1986, Ryd et al. 1993, Wilson et al. 2010). Clear limits between true migration and mere elasticity have not been established for DLRSA of other implant types or bone fractures. Moreover, as already discussed, mechanical fracture union is not a definite event. Rather, stiffness and strength of the fracture site increase with the stage of healing. To overcome this issue in interpreting DLRSA results, it has been suggested that the detected inducible micromotion can be compared to an estimated cut-off point (Wilson et al. 2010). For studies concerning fractures, it has been suggested that delineating the fracture-site micromotion with the clinical results may also facilitate the estimation of the relevant cut-off point (Chehade et al. 1997).

Table 4. Clinical RSA studies on fracture micromotion.

Year	Authors	Fracture site	Treatment modality	n	Inducible migr.	RSA follow-up
2020	Ladurner et al.	Pelvic ring	Plate and external fixator	6	Yes	2 years
2020	Galea et al.	Distal femur	Locking plate	16	Yes	1 year
2018	Bojan et al	Trochanteric	Intramedullary nail	20	No	1 year
2015	Thewlis et al.	Tibial plateau	Open reduction internal fixation	9	No	1 year
2015	van Embden et al.	Femoral neck or trochanteric	Dynamic hip screw or intramedullary nail	31	No	1 year
2014	Solomon et al.	Tibial plateau	Open reduction internal fixation	15	No	4 years
2012	Madanat et al.	Distal radius	Volar plate	15	Yes	1 year
2011	Solomon et al.	Tibial plateau	Open reduction internal fixation	7	Yes	1 year
2009	Chehade et al.	Distal femur	Locking plate	6	Yes	6 months
2008	Downing et al.	Distal radius	Volar plate	9	Yes	1 year
2004	Mattsson et al.	Trochanteric	Sliding screw with or without cement	26	No	6 months
2003	Mattsson et al.	Femoral neck	Cannulated screws with/without cement	40	No	6 weeks
2001	Kopylov et al.	Distal radius	External fixation or cement with splint	23	No	3 months
1994	Ryd et al.	Tibial plateau	Open reduction internal fixation	5	No	1 year
1993	Ragnarsson et al.	Femoral neck	Cannulated screws	16	No	Up to 1.25 years
1992	Ragnarsson et al.	Femoral neck	Cannulated screws or hook-pins	46	No	Up to 2.5 years
1991	Ragnarsson et al.	Femoral neck	Hook-pins	29	No	Up to 2.5 years
1991	Ebbinghaus et al.	Trochanteric	Sliding screw	8	No	1 year
1989	Ahl et al.	Ankle (varied)	Internal fixation	99	No	1.5 years
1989	Ragnarsson et al.	Femoral neck	Hook-pins	16	No	1 month
1986	Ahl et al.	Lateral malleolus	Internal fixation	46	No	3 months

2.5 Reporting RSA Data and Statistical Analysis

2.5.1 Reporting RSA-Measured Rigid Body Migration

Rigid body kinematics has been one of the corner stones of RSA's claim to success. As it turns out, describing rigid body motion mathematically – or indeed intuitively – is a non-trivial matter. The observed motion of bone implants is most often reported based on the principles set by Leonhard Euler (1776). In his seminal work Euler proved that any motion of a 3D rigid body could be described by translation of, and rotation around, a point in the rigid body. Because the choice of the said point in a rigid body is arbitrary there is in most cases an infinite number of ways the motion of a rigid body can be represented as scalar values. Therefore, to enable the comparison of rigid body migration between patients and RSA studies, it is essential to standardize the way this motion is reported (Valstar et al. 2005).

The obvious choice for a point of measurement is the center of gravity for a group of RSA markers. The issue with this approach is that the center of gravity will change according to the configuration of markers making comparison of results between studies and even individuals problematic. Alternatively, another well-defined feature or point in the rigid body could be used, such as the femoral head of a hip stem. In either case, standardization of the point of measurement is important to aid with repeatability and comparability of the test results. Besides the point of measurement, the choice for orientation of the coordinate system is equally critical for repeatability and should be reported (Valstar et al. 2005).

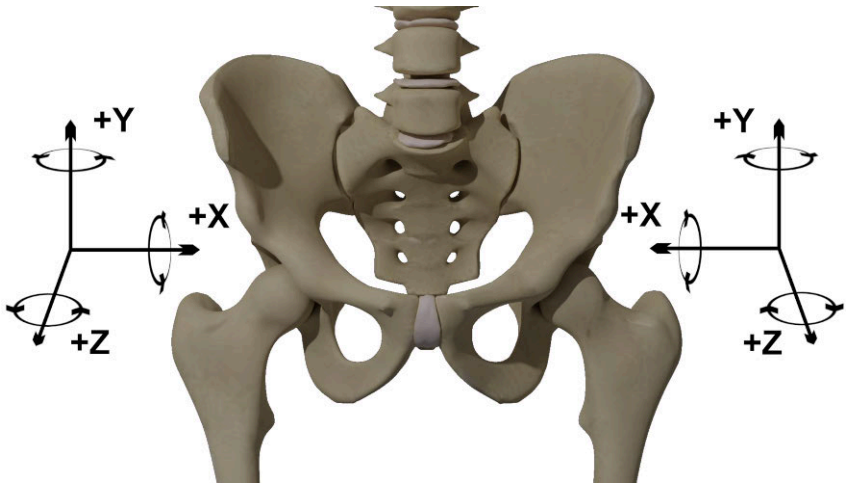


Figure 5. Orientation of the orthogonal X, Y, and Z -axes in RSA studies of the hip. Skeletal 3D model courtesy of Armen Barsegyan (cgicoffee.com).

The translation in an RSA study is often described as the components of motion in an orthogonal coordinate system with X, Y, and Z -axes (**Figure 5**). The rotation of an implant is typically presented as Euler angles (or Euler rotations) on the same orthogonal X, Y, and Z -axes. It has been stressed that the Euler angles are non-commutative and therefore should be calculated in a consistent sequence for repeatability. The RSA guidelines recommend that the Euler angles are calculated in XYZ-sequence (Valstar et al. 2005). Strictly speaking, the non-commutative nature of Euler angles degrades the comparability of the scalar rotation values both between individual patients and between separate studies. Fortunately, the small scale of rotations in a typical RSA study effectively diminishes the issues with non-commutation of Euler angles (Selvik 1989, Valstar et al. 2005).

Although intuitive, Euler angles have further mathematical limitations. For example, certain rotation sequences may result in a situation termed *gimbal-lock* (Valstar et al. 2005). This phenomenon occurs when earlier-in-sequence Euler rotations lead to a later-in-sequence rotation axis of the rigid body aligning with an axis on which the rigid body has already been previously rotated. In such a case there is a loss in degrees of freedom for rotation that can be described. This limitation can be circumvented by using more complex mathematical constructs like quaternions for describing implant rotation (Valstar et al. 2005). In studies where larger magnitudes of rotation are present (e.g., joint kinematics) the use of screw axis for describing rigid body motion has also been suggested (Valstar et al. 2005).

Even with these considerations in mind, presenting rigid body migration numerically is non-trivial. As a rigid body is free to rotate and translate about all three spatial dimensions, so do also RSA measurements have three degrees of freedom for both translation and rotation (see **Box 1** for proof). Without any of the resulting total six degrees of freedom it is not possible to definitively describe the 3D rigid body migration. This mathematical ground truth presents a problem from the viewpoint of reporting RSA results as a six-dimensional variable is inherently non-intuitive and difficult to compare between groups. Even seemingly simple concepts such as mean migration or direction of migration in a RSA study population become somewhat complex issues due to the multivariate nature of the data (Derbyshire et al. 2009).

It is no wonder then that measures for dimensionality reduction of RSA data have emerged. Maximum total point motion (MTPM) was originally introduced in the study of total knee arthroplasty as "an approximate synthesis of rotation and translation" representing "a simple [scalar] way to express the magnitude of the motion" (Ryd 1986). MTPM is by definition the largest point motion of a rigid body (Valstar et al. 2005). For obvious geometrical reasons, a combination of translation and rotation occurring simultaneously results in increased MTPM. Indeed, MTPM has facilitated the analysis and presentation of some RSA data and there is a strong precedence for the use of MTPM especially in the study of total knee arthroplasty

where rotation and translation are often seen in conjunction (Ryd 1986, Pijls et al. 2012b).

Similarly, when the magnitude of migration is of interest, the magnitude of RSA measured translation (often termed total translation [TT]) can be calculated using Pythagoras theorem to calculate the length of the translation vector (Derbyshire et al. 2009). Analogously, total rotation (TR) is often calculated from the Euler angles in RSA studies (Bontempi 2020). It should be noted that since Euler angles are not components of a single vector, this calculation does not, strictly speaking, make mathematical sense. However, it has been shown that for the small values of rotation present in a typical RSA study the errors made in calculating TR are negligible for estimating the magnitude of overall rotation (Bontempi 2020).

2.5.2 Statistical Analysis of RSA Data

As already discussed, well designed and executed clinical trials should teeter at the edge of the minimal but statistically speaking sufficient sample size. It is, then, inevitable that even relatively small errors in the statistical methodology may cause considerable errors in performing statistical inferences. The already complex issue of presenting RSA data would seem to imply that the related statistical analysis would be a similarly complicated matter. Regardless, the current guidelines for RSA studies mention statistical analysis only in passing (Valstar et al. 2005, Derbyshire et al. 2009, International Organization For Standardization 2013). While the methodological quality of RSA studies as contrasted by the published guidelines has been assessed previously, the methodology of statistical analysis in RSA literature has not been systematically evaluated (Madanat et al. 2014). Meanwhile, errors in statistical methodology have been implicated as one of the culprits for poor reproducibility in biomedical research (Dexter and Shafer 2017).

Potential Error Sources Specific to RSA

From the viewpoint of statistical analysis, the principal challenge of RSA data is the multivariate outcome measure (i.e., 3D migration) with, typically, multiple follow-ups and missing datapoints. During the early days of RSA, statistical methods for properly modeling such data structures had not yet been fully established (West et al. 2006). In the absence of multivariate analytical methods, the only available alternatives are applying univariate analytical methods to the data one degree of freedom at a time or dimensionality reduction of the outcome variable before applying univariate analysis methods (Bandyopadhyay et al. 2011). Furthermore, there is a definite barrier-of-entry related to the mathematical complexity and extensive variety of various multivariate analytical methods (Habeck and Stern

2010). Thus, it seems likely that the use of multivariate analytical methods may not have disseminated to modern RSA research.

The issue with individual univariate analyses of the RSA data, without dimensionality reduction, is twofold. First, running multiple individual statistical tests on the same outcome variable, RSA-measured migration, runs the risk of false positive results due to statistical multiplicity (Dmitrienko and D'Agostino 2018). This issue of statistical multiplicity is aggravated further if individual statistical analyzes are performed over multiple follow-ups inflating the number of statistical tests further. Even if appropriate measures are taken for correcting the familywise error rates, individual statistical test are bound to miss on multivariate phenomena in the data (Verbeke et al. 2014). As a visualization of this issue, one may consider the effective confidence intervals in the case of either a univariate or multivariate analysis of migration. In essence, the true, multivariate confidence interval of an RSA measurement forms a multidimensional ellipsoid, with likely autocorrelations among the repeated measures (**Figure 6**). Meanwhile univariate analysis has essentially a rectangular confidence interval that misses out on combined migration occurring on multiple axes simultaneously, as well as possible correlations in the data. As a combined effect, individual statistical tests have reduced statistical sensitivity and specificity (i.e., inflated type I and type II statistical error rates) in the analysis of multivariate phenomena.

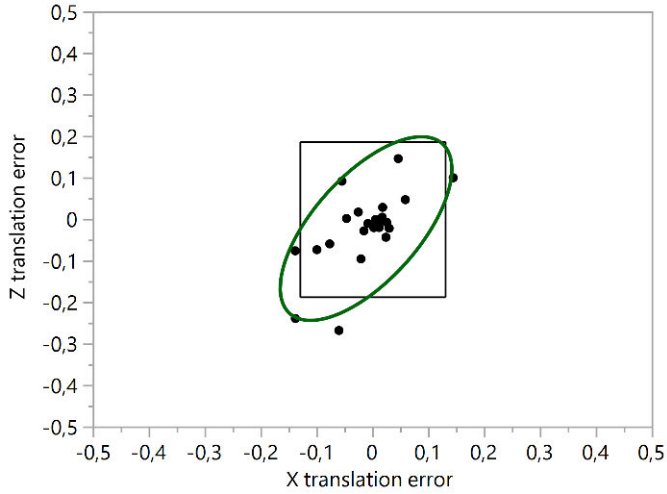


Figure 6. The measurement error of RSA double measurements plotted on a plane. In green, the 95% confidence ellipsoid of containing a double measurement (a two-dimensional representation of a multivariate confidence interval). In black, the confidence interval specified separately for each axis (95% confidence interval according to the critical value of Student's *t*-distribution with df = 24). The multidimensional confidence interval accommodates the positive correlation between the axes (autocorrelation) and has a visibly better fit on the data.

The established dimensionality reduction techniques in RSA studies are MTPM and total translation/rotation. Although, admittedly, their use may not always be motivated by the sole goal of reducing dimensionality (Valstar et al. 2005, Derbyshire et al. 2009). The principal weakness off these measures, as with any dimensionality reduction technique, is that the reduction in degrees of freedom comes with an inevitable loss of information on the true rigid body migration (see Box 1). Moreover, a critical feature of RSA – especially uniplanar RSA – is that the measurement error can vary significantly for the different axes, as discussed previously. In some specialized setups the measurement error can differ up to an order of magnitude (Garling et al. 2005). This results in situations where statistically significant migration on one axis could be overshadowed by the measurement noise of a less precise axis when MTPM, TT, or TR are calculated. The conclusion is that neither MTPM nor TT/TR can be used to definitively exclude statistically significant migration as measured by RSA. An additional concern relates to the possibility of the dimensionality reduction techniques unnecessarily inflating the number of statistical tests. If done in addition to comparing the raw rigid body migration, separately analyzing MTPM and TT/TR leads to multiple comparisons of essentially the same migration data.

Finally, these difficulties in univariate analysis of RSA data may lead to invaluable RSA data being excluded from data analysis. It is possible that data is omitted from analysis for fear of running into issues with multiplicity, for example. The fact that evidence of clinically significant migration is tied to individual axes at a fixed time point (Pijls et al. 2012b, 2012a, de Vries et al. 2014, van der Voort et al. 2015) may serve to exacerbate this issue further as valid data on secondary axes and time points are discarded.

Statistical Practices in RSA studies

Examining the RSA literature (**Table 5**), published during a two-year period from 2015 onward, it seems that the multivariate nature of RSA data is not reflected in the statistical methodology: during this period, no study used multivariate analysis methods across more than any one axis at a time. Ten studies used methods suitable for longitudinal data like repeated-measures ANOVA or LMM but still analyzed the data on an axis-by-axis basis. Twenty-seven studies used MTPM or TT/TR to facilitate the data analysis. In studies of knee joint implants, MTPM was used in eight out of 11 studies.

Table 5. Data extracted from the RSA studies, comparing two or more groups, published from 2015 until the end of 2016.

	N (%)
Total reviewed studies	42 (100%)
Migration analyzed at all follow-ups	20 (48%)
Migration analyzed on all available axes	21 (50%)
Multivariate statistical analysis (over-axes)	0 (0%)
Longitudinal statistical analysis	10 (24%)
All available RSA data analyzed	16 (38%)
Multiple individual statistical tests*	37 (88%)
– Considered or corrected for multiplicity	7 (17%)
Used TT/TR or MTPM	27 (64%)
– Analyzed both axis-by-axis data and MTPM or TT/TR	14 (31%)
All RSA migration data (axes) reported	26 (62%)

*For the same null hypothesis of "no difference in mean migration between groups"

Twenty-one studies excluded some axes of measurement completely from the data analysis. Twenty studies did not include all available timepoints into the data analysis. A total of 16 studies omitted one or more axes of measurement from the reported results. Sixteen out of the 42 studies included all available time-points and axes of measurement in the data analysis. 35 studies performed multiple individual statistical tests on the same RSA data. Seven studies clearly reported either having corrected the statistical results for multiple tests or the reason why correction was deemed unnecessary.

Analysis of Multivariate, Longitudinal Data

The choice of the ideal multivariate statistical model for RSA data could be debated at length. However, the availability of established software packages and the overall flexibility of linear mixed-effects modeling (LMM) and its generalizations have established it as one of the most popular methods for longitudinal and multivariate longitudinal data (Verbeke et al. 2014). Various categories and algorithmic variations of mixed-effects models exist (Jiang 2007). Of these methods, Gaussian LMMs have been extensively studied and utilized in the analysis of real-valued longitudinal data such as that produced by RSA studies (Asar et al. 2020).

The first known formulations of linear random-effects models date back the monography of astronomer George Biddell Airy published in 1861 (West et al.

2006). Airy's random-effect model had only a single random effect that he used to model the variance that was caused by making telescopic sightings at different occasions. Still, it was not until the 1980s that the method was successfully applied to analysis of longitudinal data. Thereafter, LMM started seeing increasing application in medical sciences from early 1990s onward (West et al. 2006). Initial enthusiasm was focused on univariate applications, but LMM was eventually established in the analysis of multivariate, longitudinal data in the 2000s (Verbeke et al. 2014).

Perhaps the main advantage of LMM, especially from the viewpoint of RSA, is that it allows for unbalanced data (Cnaan et al. 1997, Peters et al. 2012). This means that the number of observations need not be equal across subjects, time, or axes of migration. Missing observations leading to unbalanced data are exceedingly common among RSA studies (see, for example, **Table 2**). As a comparison, a method that relies on a balanced dataset, such as multivariate analysis of variance (MANOVA), would require either exclusion of entire study subjects with unbalanced data or imputation of missing observations (Finch 2016). Consequently, use of MANOVA has been discouraged in such datasets in favor of mixed-effects models (Armstrong 2017). Furthermore, LMM has relatively lax assumptions on the structure of the underlying data. For example, compared to MANOVA, LMM does not assume sphericity of data and the covariance-variance matrix across different levels of the repeated-measures is not constrained to any specific pattern (Armstrong 2017). For truly multivariate data – like RSA migration – a structured covariance-variance matrix is unlikely (Verbeke and Molenberghs 2012).

It has been stressed that having multivariate outcome measures at hand does not necessarily warrant multivariate analytical methods (Verbeke et al. 2014). Such a situation arises when multivariate outcomes are not of interest from the viewpoint of the research question. For RSA research, it is rather easy to make the case that multivariate outcomes are of equal interest to univariate outcomes. First, even when properly standardized, the choice for the orientation of the reference coordinate system in an RSA study is somewhat arbitrary (Valstar et al. 2005). Therefore, it is unreasonable to assume that statistically or, indeed, clinically significant 3D migration would be constrained to any single axis. Secondly, it could be argued that the value of detecting multivariate migration (e.g. rotation and translation of an implant occurring simultaneously) is already a recognized fact in the study of total knee arthroplasty as demonstrated by the relative success of MTPM (Ryd 1986, Pijls et al. 2012b). Lastly, as discussed previously, it is possible that the migration pattern (for definition, see **Box 4**) varies across various strata of implant design, patient related factors, and even surgical considerations (de Vries et al. 2014). Therefore, the possibility of unexpected migration patterns should not be dismissed in most RSA study settings.

BOX 4. The migration pattern

De Vries et al. (2014) defined the *migration pattern* of an implant as whether migration or an aspect of it was expected or not. In the context of multivariate longitudinal analytical methods, a more general definition seems logical: the presentation of migration as a function of both time and the measurement axes. This broader definition is intended in references to *migration pattern* in this thesis.

Limitations of LMM come in the form of the statistical assumptions necessary for model fitting and statistical inference. For Gaussian LMM these assumptions include a linear relationship between the predictors and the response, random occurrence of missing data ("missing at random" (Rubin 1976)), normality of model residuals, homoscedasticity of model residual variances, no multicollinearity of predictors, normality of random effects, and independence of subjects (Schielzeth et al. 2020). Remarkably, LMM is quite robust to even significant violations of its key assumptions (Schielzeth et al. 2020). A noteworthy consequence of the distributional assumptions is that Gaussian LMM is susceptible to outlier introduced bias. This is a shared property for most parametric analytical methods and, contrary to a common misconception, not even non-parametric methods are entirely immune to outliers (Zimmerman 1994). Fortunately, the core model assumptions of LMM including the distributional assumptions can be evaluated with, for example, graphical methods (Singer et al. 2017). Further, influence diagnostics can be used to evaluate the impact a single observation or a group of observations has on the resulting model thereby helping in the detection of significant outliers (Zewotir and Galpin 2005).

3 Aims

The general aim of this thesis was to study RSA methodology and its application in research of cementless total hip arthroplasty in hip osteoarthritis patients as well as internal fixation of femoral neck fractures.

The individual studies of this thesis were aimed specifically:

1. to study the application of marker-based RSA technique for evaluation of a patient-related factor, systemic bone mineral density, as a potential determinant of primary ABG II cementless cup stability among female patients.
2. to study whether the accuracy and clinical precision of model-based RSA, using computer-aided design models, are comparable to that of marker-based RSA in cementless total hip arthroplasty using Accolade II cementless femoral stem.
3. to evaluate the feasibility of marker-based differentially loaded RSA for the assessment of the fracture-site stability and inducible micromotion in patients with internally fixed femoral neck fractures.
4. to examine if multivariate linear mixed-effects model is suitable for comprehensive 3D analysis of RSA-measured migration of Accolade II cementless femoral stem.

4 Hypotheses

The study specific hypotheses were:

1. Marker-based RSA can be used for delineating the impact of systemic bone mineral density on the early two-year micromotion of custom-designed cementless acetabular cups.
2. Model-based RSA has comparable accuracy and precision to conventional RSA in the study of a cementless parallel-sided tapered wedge femoral stem.
3. Differentially loaded RSA can be used to monitor femoral neck fracture healing *in vivo*.
4. Multivariate linear mixed-effects model can be used to compare intergroup 3D migration of cementless femoral stems in the clinical setting possibly improving statistical power.

5 Materials and Methods

Table 6. The key demographics of all clinical studies in the thesis. Qualitative numbers given as median (range).

	Study I	Study III	Study IV
n (female/male)	34 (34/0)	16 (10/6)	57 (57/0)
Age (years)	65 (41 to 78)	74 (56 to 90)	68 (60 to 84)
– **Age ≥ 75 (n)**	1	8	9
BMI (kg/m^2)	29 (21 to 48)	23 (16 to 28)	27 (16 to 43)
BMD			
– **Lowest T-score**	-1.8 (0.3 to -3.7)	-2.6 (-0.5 to -4.8)	-1.2 (-3.6 to 1.1)
– **T-Score ≤ - 1 (n)**	24	15	30

5.1 Study I: ABG II Acetabular Cup

This study examined the two-year RSA measured migration of a cementless acetabular cup in a prospective cohort of consecutive female patients who underwent cementless THA due to advanced osteoarthritis. Specifically, the study was designed to delineate the effect the patient-related factor of systematic BMD had on the observed RSA-measured migration. The results for the femoral stem in this cohort have been described previously (Aro et al. 2012).

Patients and Surgery

The study sample was recruited from a cohort of 110 consecutive postmenopausal females who were scheduled for THA due to osteoarthritis at Turku University Hospital. The recruitment window was open August 2003 until March 2005. The main inclusion criteria were (1) a female patient, (2) age of less than 80 years, and (3) the American Society of Anesthesiologists (ASA) classification of less than or equal to three. During the recruitment 61 patients consented to screening. The exclusion criteria were aimed at minimizing confounding factors as to the patients systemic and local skeletal status: (1) inflammatory arthritis, (2) untreated parathyroid condition, (3) past or present osteoporosis-medication, (4) ongoing

corticosteroid therapy, (5) any other medication affecting bone metabolism, or (6) severe osteoporosis (necessitating intervention). After screening 43 patients were enrolled into the study with further three dropouts due to early periprosthetic fracture and six patients due to poor visualization of RSA markers. Thus, 34 patients (**Table 6**) remained for the data analysis.

The surgeries were done through the anterolateral Hardinge approach and the initial stability was achieved by under-reaming and then press-fitting the acetabular cup. All patients received a custom RSA-marked cementless ABG II (Anatomic Benoist Girard II, Stryker Europe) THA implant with ceramic-on-ceramic bearings. Four to six tantalum markers were inserted into the cancellous bone surrounding the acetabular socket to serve as the RSA reference rigid-body for the acetabular cup. The acetabular cup had eight factory-applied tantalum beads at the cup opening. Five pegs with tantalum beads at their tips were used for augmenting the fixation with the secondary goal of further facilitating the RSA analysis.

Systemic Bone Mineral Density

The aim of the study was delineating the effect a patient related factor, i.e. systemic BMD, might have on the RSA-measured migration. To achieve this goal, the preoperative systemic BMD was assessed for each patient using dual-energy X-ray absorptiometry (DXA). The DXA measurement was performed at three sites – proximal femurs, lumbar spine, and the distal radius of the non-dominant hand. To facilitate the data analysis, the lowest T-score for any of these sites were used to stratify the patients into two groups: patients with a T-score less than or equal to -1 were inducted in the “low BMD” -group and the remaining patients were considered to have normal BMD.

RSA

The translations and rotations of the acetabular cups were analyzed using marker-based RSA at three, six, 12, and 24 months after surgery as compared to baseline examination within 7 days post operatively. The accuracy of the RSA setup has been evaluated previously using a phantom model (Mäkinen et al. 2004). Before considering any RSA measurement valid, the stability of individual markers constituting the relevant rigid-bodies was assessed using ME and markers with ME exceeding 0.35 were excluded. Similarly, the sufficient spatial distribution of the markers was confirmed using CN and rigid-bodies above the limit of 150 were considered unacceptable. The clinical precision was confirmed through double measurements of zero-motion. All RSA analyses were performed using UmRSA version 6.0.3.7 (RSA BioMedical Innovations AB, Umeå, Sweden).

Clinical Outcome Assessment

The clinical outcome of the surgery was followed using Western Ontario and McMaster Universities Osteoarthritis index (WOMAC) and Harris Hip Score (HHS) as measures of patient reported and physician reported outcome, respectively. All patients were followed with conventional radiography up to two-years with an additional late follow-up at a median of 8 years (range 2-10 years).

Statistical Analysis

The association between the assigned BMD group and the RSA-measured migration at two years was analyzed as the main outcome. Further, the study groups were assigned a risk-of-revision classification according to the observed proximal migration (Pijls et al. 2012a). Univariate statistical analyses were performed on an axis-by-axis basis. Mann-Whitney U-test was used when the assumption of normality was breached. A two-sample t-test was preferred whenever possible. Further, in an attempt to control for potential confounders, a *post hoc* linear regression controlled for age and body mass index (BMI) was decided on for any statistically significant outcome in the main outcome measure. Repeated-measures ANOVA and Friedman test were used to analyze the migration data in either group for statistically significant migration beyond the settling period of three months. All analyses were corrected for family-wise multiplicity using the Bonferroni correction. All statistical analyses were performed using SPSS for Windows version 21.0 (IBM SPSS Statistics, Armonk, NY).

5.2 Study II: MBRSA of the Accolade II Stem

This sub-study examined whether the core advantage of RSA – high accuracy – holds when model-based RSA is applied in lieu of marker-based RSA using a parallel-sided tapered-wedge femoral stem. The main part of the study was confirming the *in vitro* accuracy of model-based RSA using a phantom model. The clinical precision of the method was also confirmed in a cohort of patients participating in a subsequent clinical trial.

Phantom Setup

The accuracy over a range of likely migration ranges for a femoral stem was assessed using a phantom model (**Figure 7**) involving the studied femoral stem (Accolade II, Stryker Orthopaedics, Mahwah, NJ, USA). In the phantom model, the complete THA system, including the femoral stem, the cobalt-chrome femoral head, and the relevant titanium cup with a polyethylene insert, was fixed securely on a rigid base plate in the anatomic position. In the final phantom model, a transparent plastic tube

was used as the phantom for the proximal femur. This allowed directly observing that the femoral stem was not in contact with the femur phantom and unobstructed manipulation of the phantom was assured. The simulated migrations in the phantom were achieved by attaching the femur phantom to a high-precision manipulator with a translation stage on X-, Y-, and Z-axes and a rotation stage on the Y-axis. To enable a direct comparison of the model-based and marker-based methods, three tantalum markers on plastic pegs were glued onto the femoral stem using the typical positions in clinical RSA trials. Subsequently, the phantom was imaged repeatedly with displacement introduced systematically on each axis available for manipulation.

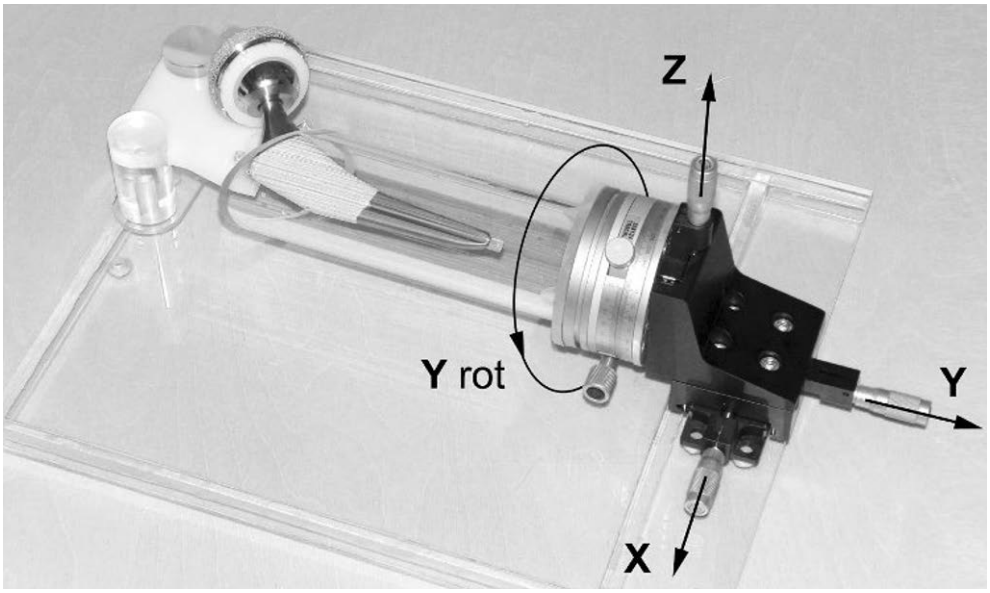


Figure 7. The phantom setup of the Accolade II stem. An RSA-marked plastic tube representing the proximal femur is attached to a four-axis high-precision manipulator for the introduction of controlled micromotion. The RSA-marked femoral stem is attached rigidly to the manipulator via the baseplate.

Clinical Cohort

To delineate how the *in vitro* results transferred to the clinical setting, the clinical precision of marker-based RSA was also assessed in a cohort of 24 patients using double measurements (Aro et al. 2019).

RSA

The stereoradiographs of the phantom were analyzed with two separate software suites. MBRSA software (Version 3.34; Medis Specials BV, Leiden, The

Netherlands) and UmRSA version 6.0.3.7 (RSA BioMedical Innovations AB, Umeå, Sweden). The MBRSA software was used to perform both the model-based analysis and a comparative marker-based analysis. UmRSA was utilized to make a second marker-based analysis to confirm that the results were comparable between the two software packages. The analysis of the clinical cohort was done using the MBRSA software, only, as the implanted stems had no RSA markers necessitating a model-based analysis. A combined head-stem model was used for the model-based analysis (Prins et al. 2008). For the model-based analysis of the stem, manufacturer-supplied computer-assisted design (CAD) surface models were utilized.

Statistical Analysis

The *in vitro* accuracy and clinical precision were calculated according to established principles (Ranstam et al. 2000). All analyses were performed using SPSS for Windows version 23 (IBM SPSS Statistics, Armonk, NY).

5.3 Study III: DLRSA of Internally Fixated Femoral Neck Fractures

The third study was a prospective cohort study on patients considered for internal fixation due to an isolated fracture of the femoral neck at the Turku University Hospital.

Patients and Surgery

Recruitment occurred from September until June 2011. The inclusion criteria were (1) age over 50 years with (2) a fracture type suitable for internal fixation (AO/OTA fracture types 31-B1, 31-B2, and 31-B3). The exclusion criteria were factors considered as contra-indication for internal fixation (e.g., diagnosis of rheumatoid arthritis, severe osteoarthritis, pathological fracture, or corticosteroid use). Surgical treatment of each patient was scheduled within 48 hours of arriving at the hospital. During the recruitment period between September 2010 and June 2011, a total of 321 patients presented with a fracture of the femoral neck at Turku University Hospital. Seventeen patients were assessed for eligibility and ultimately 16 patients (**Table 6**) were enrolled into the study.

The original study protocol included a randomization of patients to internal fixation with either cannulated screws or a sliding hip screw. Reliable insertion of multiple markers into the femoral head proved difficult in the case of the dynamic hip screw and the randomization was stopped after two patients had been operated with the sliding hip screw. All subsequent patients were operated with cannulated

screws. During surgery, three or more markers were inserted to the femoral head through the holes drilled for the screws. Further, a minimum of three markers were placed in the region of greater trochanter to serve as the reference rigid body for the femoral head.

RSA and DLRSA

Marker-based RSA was used to measure the sustained translations and rotations of rigid body formed by the femoral head markers in relation to the rigid body formed by markers in the greater trochanter. Baseline RSA imaging was performed within three days after surgery with follow-ups at six, 12, 24, and 52 weeks. Additionally, DL-RSA was performed at each follow-up including at baseline. Loading of the hip was achieved by the patient pressing the operated foot against a force plate with voluntarily maximal force while providing a counterforce with both hands. The second RSA imaging was performed during the measured peak loading. Both unloaded and loaded imaging were performed in the supine position. The detected migration between each unloaded follow-up and baseline imaging was considered as sustained micromotion. The migration between the loaded and unloaded conditions at each follow-up was considered as inducible micromotion. Fracture site compliance was calculated as the inducible total translation or rotation divided by double the force measured at the force plate (Denham 1959).

Analogous to the first study (I) ME and CN were used to evaluate the stability and spatial distribution of the markers. Due to the physical constraints of the femoral head somewhat higher condition numbers of up to 244 were accepted. Similarly, the maximum accepted ME was 0.58. Three patients had to be excluded from the RSA analysis due to an insufficient number of stable markers. The precision of measurements both with and without loading were confirmed through double measurements for each patient. All RSA analyses were performed using UmRSA version 6.0.3.7 (RSA BioMedical Innovations AB, Umeå, Sweden).

Conventional Radiography

The baseline anteroposterior radiograph was taken immediately post operatively whereas baseline RSA was performed within three days post operatively. Beyond this crucial difference, conventional radiography was done during the same follow-up visits as RSA. The reduction of the femoral abductor moment arm (offset) and overall femoral length (shortening) were measured from anteroposterior radiographs at each follow-up beyond baseline. Additionally, these measurements were projected onto the femoral neck shaft to arrive at a measurement of femoral neck shortening

(FNS). Each radiograph was calibrated individually using the known thickness of the cannulated or locking screws at the fracture site.

The plain radiograph measurements were facilitated by a custom-programmed software that was developed by the author for this sub-study. The software was based on the methodology described by Zlowodzki et al. (2008). However, Zlowodzki et al. used the radiograph of the unoperated side as the baseline. Since interfragmentary micromotion was of main interest in the present study, instead of comparisons with the unoperated side, the current analysis was performed using radiographs of the operated side only. This also allowed the superimposition of the calibrated radiographs between examinations for a more exact match.

The reliability of the measurement methodology and implemented software was assessed by calculating intra- and interobserver agreement on repeat measurements of same image pairs. Additionally, the accuracy of the measurements was evaluated using RSA as the gold standard. Due to the differences in timing the baseline imaging, comparison of RSA and conventional radiography was done only for changes in migration beyond the six-week follow-up. RSA measurements of X-translation, Y-translation and total translation were compared with the measurements of offset, shortening and FNS, respectively.

The development of fracture union was assessed from standard radiographs where fracture union was defined as the disappearance of fracture lines in the cortical and trabecular bone in an asymptomatic patient (Bhandari et al. 2013).

Statistical Analysis

Intra- and interobserver agreement was evaluated by calculating intraclass correlation coefficients for the related observations. Bland-Altman-plots and the related 95% limits of agreement (LOA) were used to compare conventional radiography with RSA. The association between inducible and sustained micromotion was delineated using linear regression. The 95% confidence intervals for zero-motion (0.4 mm for total translation and 1.6 degrees for total rotation) were used as limits of detection for inducible micromotion at the individual level. The Bland-Altman-plots and related statistics were generated using R version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria). All remaining analyses were performed using SPSS for Windows version 25 (IBM SPSS Statistics, Armonk, NY).

5.4 Study IV: Multivariate LMM Analysis of Accolade II stem

The fourth study used previously published clinical data from an RSA trial to describe the process of comparing RSA measured micromotion between two groups using multivariate linear mixed-effects modeling.

Patients, Surgery, RSA, and Local BMD

The original trial was designed as a randomized, double-blinded, and placebo-controlled study of the effects of an antiresorptive drug, denosumab, on the stability of a cementless femoral stem (Accolade II, Stryker Orthopaedics, Mahwah, NJ, USA) in postmenopausal females. This trial was a direct continuation of the second sub study and its clinical patient cohort. The study protocol, demographics (**Table 6**), inclusion/exclusion criteria and univariate data-analyses of the original trial have been described in detail previously (Aro et al. 2019). The present multivariate LMM reanalysis was focused on the previously reported model-based RSA results of the placebo and intervention study groups. In the intervention group patients were initiated on denosumab one month before surgery with a follow-up dose at six months. Beyond baseline, RSA was utilized for the measurement of sustained micromotion in either study group at 12, 22, and 48 weeks. DXA was used to determine periprosthetic bone mineral density at the seven frontal Gruen zones post-operatively and at each subsequent follow-up (Gruen et al. 1979).

Statistical Analysis

The main outcome measure was the multivariate RSA-measured migration that was compared between the study groups using multivariate LMM on all available time-points. The model statement included both the axis of migration and the follow-up point as the repeated factors. This allowed the comparison of all axes of migration and their time-related changes between the groups using only a single statistical model. The residual variance-covariance matrices for both repeated measures were assumed as unstructured. Graphical methods were used to confirm appropriate model fit and the suitability of the linear model. To underline the limitations of dimensionality reduction techniques, MTPM, total translation, and total rotation were analyzed using LMM and the results were compared with the multivariate analysis. Previous analyses of the data have revealed influential outliers in the data (Aro et al. 2020, Nazari-Farsani et al. 2021). The previous methodology for detecting outliers focused on only Y-axis rotation and translation. In the present multivariate analysis on additional outlier was detected with nearly an order of magnitude greater X-rotation (4.32°) compared to any other patient – this patient was also excluded

from the reanalysis to facilitate a satisfactory model fit. The analyses on the clinical data were performed using SAS for Windows version 9.4 (SAS Institute, Cary, NC, USA).

Monte Carlo Simulation

The multivariate model and the related complexities of mixed effects do not necessarily make intuitive sense leading to the possibility of misconceptions. Therefore, a Monte Carlo simulation evaluating the statistical power of multivariate LMM and univariate Welch's t-test was performed on simulated RSA data. The variance-covariance structure was modeled using the clinical RSA data at the 12-week follow-up. For the analysis, two groups (n=30) were sampled from populations with identical variance-covariance parameters on six degrees of freedom (X-, Y-, and Z-rotation and -translation axes). However, a mean intergroup difference was introduced to Y-translation matching exactly the level necessary for a univariate t-test to achieve 80% statistical power on that axis at an alpha level of 0.05. Subsequently, the simulated data was analyzed with three different analytical methods: (1) a univariate Welch's t-test on the Y-translation axis, (2) univariate Welch's t-test of both total translation and total rotation, and (3) a multivariate linear LMM on all axes simultaneously. The simulation and analysis steps were repeated 3 000 times to estimate the true statistical power of each statistical test at an alpha-level of 0.05. To further underline the effect of misalignment between clinically significant migration and any arbitrary RSA axis, the described Monte Carlo simulation was repeated by offsetting the significant migration from the Y-axis by a known amount. The simulation was run at five-degree offset angle intervals from 0 to 90 degrees and the observed statistical power was graphed as a function of the offset angle.

As a demonstration of a worst-case scenario performance for the multivariate LMM, a second Monte Carlo simulation with 10 000 iterations was performed with zero covariance between all axes.

To confirm the empirical alpha levels (false positive rates) for each analytical method, a third Monte Carlo simulation was executed where the sampling populations had zero mean difference on all axes. This simulation was run for 10 000 iterations. To further demonstrate the practical implications of statistical multiplicity, multiple individual t-tests on each axis separately were included into this latter Monte Carlo simulation. If any of the multiple individual t-tests had p-value less than the alpha-level of 0.05 the result was considered false positive.

The Monte Carlo simulations were performed using R version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

5.5 Ethical Considerations

Each of the studies contained in this thesis were conducted according to the Declaration of Helsinki. The Ethics Committee of the Hospital District of Southwest Finland approved the associated study protocols. Written informed consent was required from all study participants of the clinical studies.

6 Results

6.1 Study I: ABG II Acetabular Cup

The acetabular cups in the low BMD group showed greater proximal migration (**Figure 8**). The intergroup difference was statistically significant (mean difference 0.21 mm, 95% CI: 0.06-0.37 mm, $p = 0.04$). This relationship remained statistically significant ($\beta = 0.60$, $p = 0.001$, overall $R^2 = 0.39$) even after controlling for age and BMI using multiple linear regression. Upon graphical inspection the proximal translation of the cups in the low BMD group seemed to increase even after the settling period of 3 months. This observation was confirmed by repeated-measures testing. A *post hoc* paired-samples t-test demonstrated a statistically significant increase in proximal translation between 3 and 12 months in the low BMD group. In the normal BMD group, there were no statistically significant time-related changes in proximal migration beyond the three-month follow-up (repeated-measures ANOVA, $p = 0.9$). Based on the risk classification for acetabular cups by Pijls et al. (2012a), the low BMD group as a whole was within the category of "at risk of revision" while the normal BMD group belonged to the category of "at low risk of revision". WOMAC, HHS, or the radiographic follow-up did not show a statistically significant difference between the two groups.

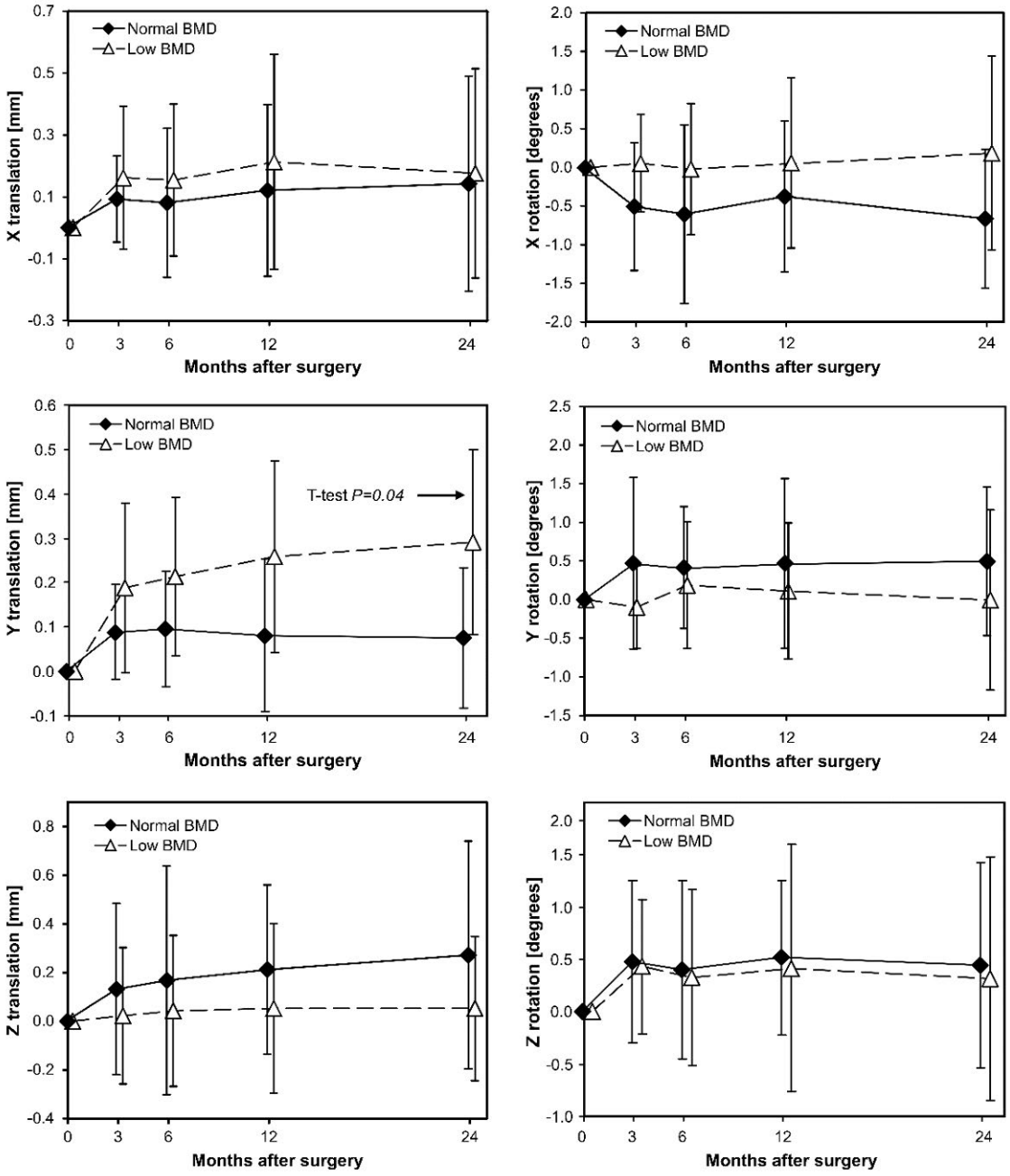


Figure 8. The migration of the ABG II acetabular cups over the whole study period factored by the assigned BMD group. Analyzed at 24 months, Y-translation showed a statistically significant difference between the two groups.

6.2 Study II: MBRSA of the Accolade II Stem

The model-based and marker-based RSA showed similar levels of measurement error and mainly negligible bias (**Table 7**). The error for determining Y-axis rotation seemed to have a slight positive bias for all methods. There were no significant differences in bias between the methods suggesting that the bias was not related to the use of the model-based method. Assuming negligible bias, the accuracy of the model-based method (centered 95% LOA) was 0.02, 0.04, and 0.16 mm for X-, Y-, and Z-translation, respectively. The accuracy for Y-rotation was 0.47 degrees (**Table 8**). The observed clinical precision demonstrated that the precision of the method was good in the clinical setting (**Table 9**).

Table 7. Bias and 95% LOA of RSA measurements made on the phantom with known displacements on each axis of interest. Translations in mm, rotations in degrees.

	Model-based RSA		Marker-based (MBRSA)		Marker-based (UmRSA)	
Axis	**Bias (95% CI)**	**95% LOA**	**Bias (95% CI)**	**95% LOA**	**Bias (95% CI)**	**95% LOA**
X trans.	0.01 (0.01 to 0.02)	-0.01 to 0.03	0.00 (-0.01 to 0.01)	-0.02 to 0.02	0.01 (-0.00 to 0.02)	-0.02 to 0.03
Y trans.	0.01 (-0.00 to 0.03)	-0.03 to 0.05	0.01 (-0.01 to 0.02)	-0.04 to 0.05	0.01 (0.00 to 0.01)	-0.01 to 0.02
Z trans.	-0.05 (0.00 to -0.11	-0.22 to 0.10	0.00 (-0.04 to 0.04)	-0.10 to 0.10	0.02 (-0.00 to 0.04)	-0.04 to 0.07
Y rot.	0.23 (0.06 to 0.41)	-0.24 to 0.71	0.15 (0.07 to 0.23)	-0.06 to 0.37	0.04 (-0.10 to 0.17)	-0.33 to 0.40

Table 8. Accuracy of the RSA measurements under the assumption of no systematic bias. Translations in mm, rotations in degrees.

Axis	Model-based RSA	Marker-based (MBRSA)	Marker-based (UmRSA)
X trans.	0.02	0.02	0.03
Y trans.	0.04	0.05	0.01
Z trans.	0.16	0.10	0.05
Y rot.	0.47	0.22	0.37

Table 9. Precision of model-based RSA. Precision is defined as the 95% confidence interval of containing a double measurement.

Axis	Precision
X trans.	0.13
Y trans.	0.14
Z trans.	0.47
X rot.	0.41
Y rot.	0.79
Z rot.	0.39

6.3 Study III: DLRSA of Internally Fixated Femoral Neck Fractures

The loading of the fracture-site induced instant micromotion that could be detected with DLRSA (**Figure 9**). Some fractures were stable, as measured by DLRSA, already at baseline and the proportion of stable fractures increased with post-operative time. Patients that achieved fracture union were characterized by no detectable inducible micromotion or inducible micromotion only at baseline and generally low fracture-site compliance (**Figure 9** and **Figure 10**). Of the five failure cases (osteonecrosis, non-union, union with malrotation), all but one displayed inducible micromotion beyond baseline. This association was confirmed as statistically significant with a *post hoc* analysis ($p = 0.029$, Fisher's exact test). Similarly, failure cases – save the late osteonecrosis case – presented as outliers in the magnitude of permanent displacement (**Figure 11**). The magnitude of post-operative inducible micromotion did not show a statistically significant correlation with sustained micromotion.

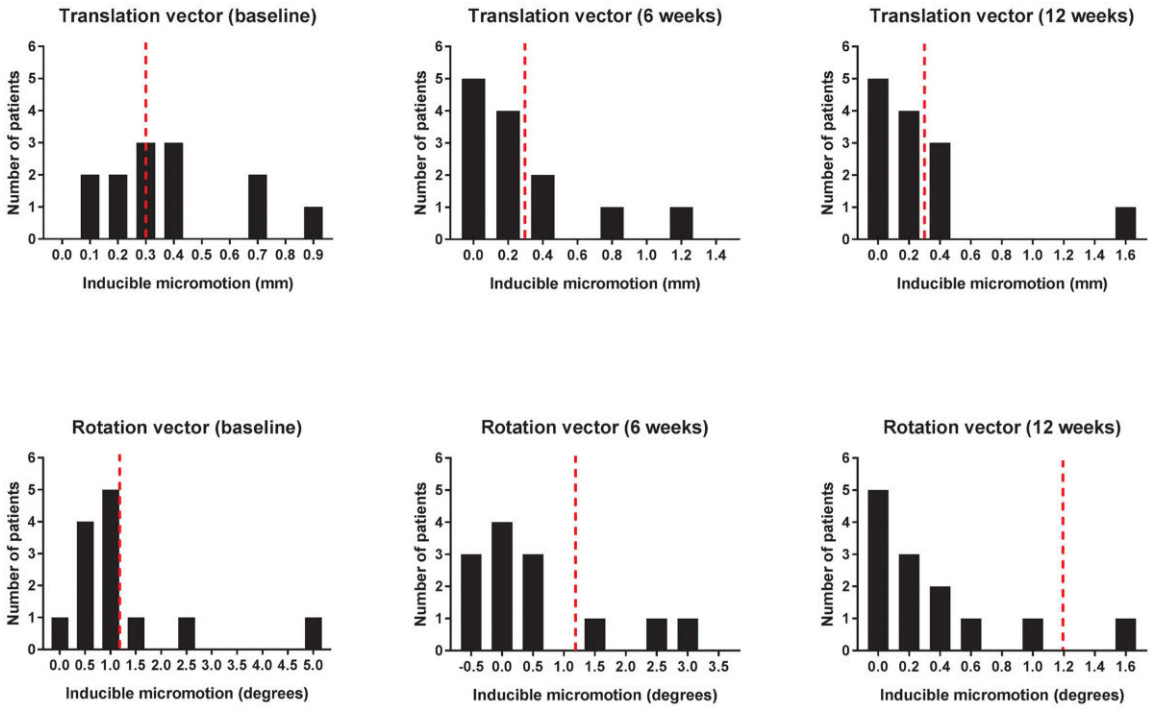


Figure 9. The inducible displacement of patients in relation to the precision limits of the DLRSA as a function of the postoperative time.

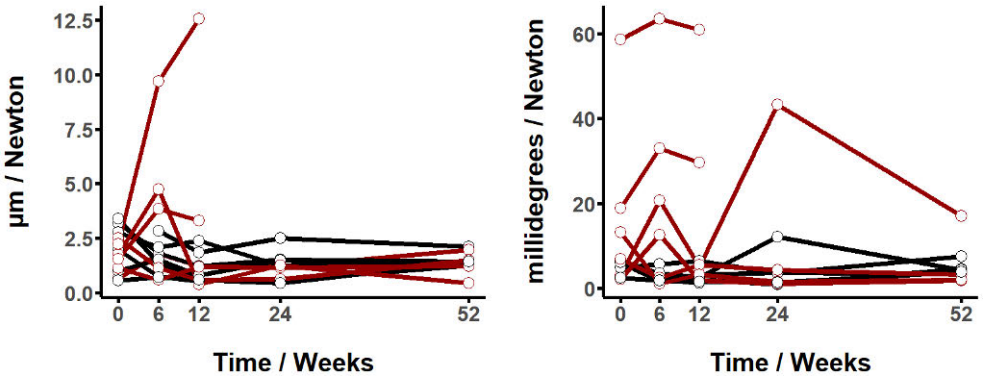


Figure 10. Translational and rotational fracture-site compliance of each individual femoral neck fracture. Clinical unions in black, clinical non-unions (including one case of union with malrotation) in red.

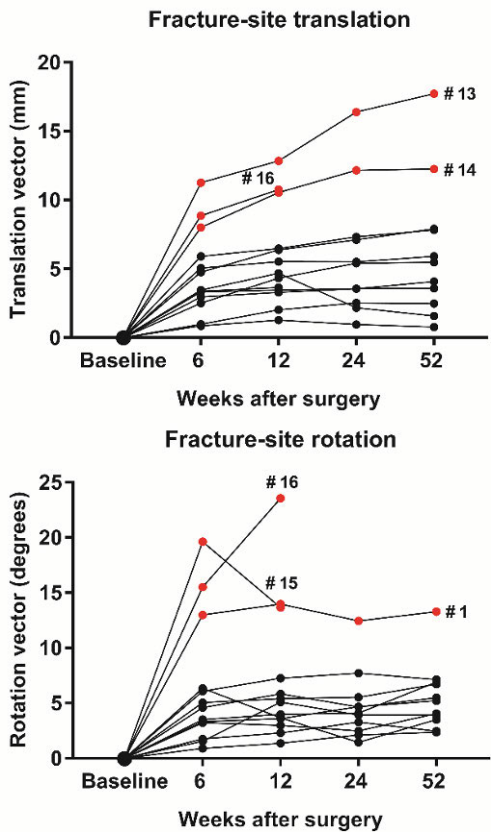


Figure 11. Permanent displacement of each proximal fracture fragment in relation to the distal femoral fragment. Outliers, all clinical failures, in red.

According to the measurements made from conventional radiographs at 52 weeks, the mean offset and shortening were 3.0 and 9.8 mm, respectively. The mean FNS along the neck shaft was 9.2 mm. The intra- and interobserver agreement on the measurements made from conventional radiographs were excellent. Intraobserver ICC was 0.99 for FNS and 0.98 for both offset and shortening. Interoberver ICC was 0.95 for FNS, 0.95 for offset and 0.91 for femoral shortening.

6.4 Study IV: Multivariate LMM Analysis of Accolade II stem

The multivariate LMM showed a trending difference in 3D migration between the study groups overall during the follow-up period ($F_{6,111} = 1.9$, $p = 0.09$). None of the dimensionality reduction measures (i.e., TT, TR, or MTPM) showed statistically significant differences between the groups (**Figure 12**). Upon inspection of the LMM parameters the main difference between the groups seemed to align with the

Z-rotation and Y-translation axes. The trending statistical result was viewed as warranting *post hoc* analyses. An inspection of the effect slices revealed a statistically significant difference in Z-rotation between the two groups could be observed ($F_{1,106}$ = 4.9, p = 0.03) with the intervention group showing greater Z-rotation. Y-axis had a trending result ($F_{1,143}$ =3.3, p = 0.07) suggesting that the 3D difference between the groups was a combination of Z-rotation and Y-translation. In the combined study population Z-rotation and Y-translation showed a statistically significant correlation (R = -0.449, p = 0.001).

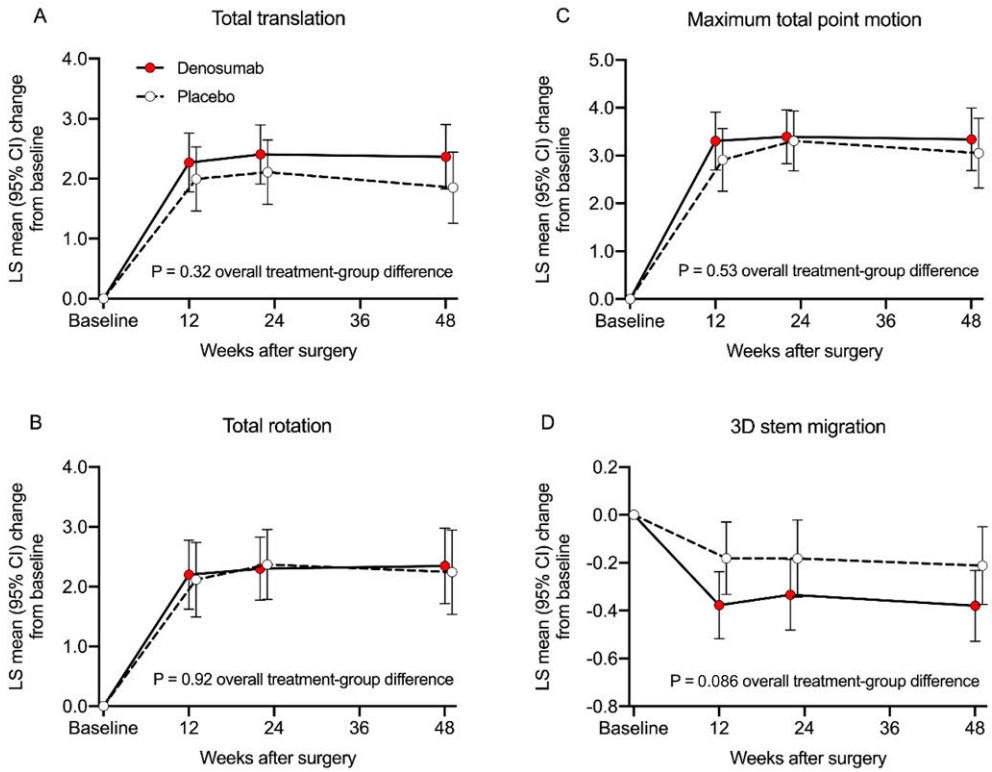


Figure 12. Least square mean estimates (95 % CI) of the estimated linear model for (A) total translation, (B) total rotation, (C) MTPM, and (D) multivariate LMM. Please note, that the units and their magnitude between the multivariate and univariate models are not comparable.

In the analysis of the simulated RSA data, multivariate LMM had similar or greater statistical power compared to alternatives (**Figure 13**). The mean statistical power of multivariate LMM was 80.6% over the whole range of offset angles. In the case of the univariate t-test and total translation/rotation the overall mean statistical power was 64.3% and 24.7%, respectively. Applied to the data with zero covariance

between the measurement axes, multivariate LMM had a worst-case scenario statistical power of 48.1%.

When comparing data with zero population mean difference, the alpha-level (false positive rate) was maintained at or below 5% for both the univariate t-test and multivariate LMM. The analysis of total translation and rotation consisting of two separate t-tests showed an alpha-level of 9.5%. When univariate t-tests were performed on all axes simultaneously the alpha-level was inflated considerably at 23.6%.

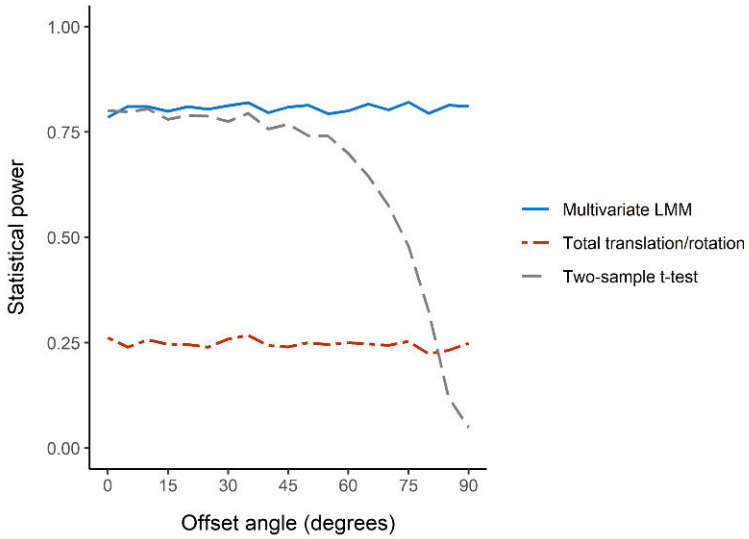


Figure 13. Empirically derived statistical power of multivariate LMM, univariate t-test, and total rotation and translation in multivariate normal data with six orthogonal degrees of freedom. The t-test has been done only on one degree of freedom with an induced difference in migration. The results are graphed as a function of offsetting the statistically significant migration away from the degree of freedom that the t-test was performed on.

7 Discussion

7.1 Study I: ABG II Acetabular Cup

RSA has been used and validated extensively for predicting the long-term outcome of THA components – like the acetabular cup – as a function of various implant related factors (Pijls et al. 2012a). The host related determinants for the clinical outcome have seen much less research focus. This issue is accentuated with cementless THA components where osseointegration is aspired. Stable initial fixation is a key factor for osseointegration, and it is best achieved with good host bone quality. Yet, quantified BMD has not previously been correlated with early acetabular cup stability. The first study of the thesis was designed to explore whether preoperative systemic bone mineral density would affect the early RSA-measured stability of the acetabular cups.

The results of our study suggest that low systemic BMD was an important factor for the early stability of the cementless acetabular cup in our study population. A closer look at the migration data suggested that the stabilization of the acetabular cups was seemingly delayed in the low BMD group. This observation fits the hypothesis that low BMD compromises the primary stability of the cups which in turn inhibits or at least delays osseointegration. As a finding of potential clinical relevance, the patients in the low BMD group were suggested to be at a greater risk of revision based on the extent of proximal migration (Pijls et al. 2012a).

Most importantly, our findings implicate preoperative BMD as a potential confounder in observational RSA studies of cementless acetabular cups – at least considering female patients. In other words, observational and even comparative RSA studies could benefit from determining and controlling for the preoperative BMD to reduce unaccounted for variation in the clinical RSA data. Further, the identification of preoperative BMD as a potential patient-related factor for implant stability and even risk of revision provides a potential target for intervention. However, the first step should still be to see if and how well our results generalize outside our target population of female osteoarthritis patients and beyond cementless fixation, only.

At the time the results of the current study were published the closest comparable research had been conducted by Digas et al. (2004). In line with the present study,

Digas et al. reported increased acetabular cup migration in patients who had a diagnosis associated with compromised bone quality. After the publication of our results Tabori-Jensen et al. (2020) have conducted a trial on the effect of both systemic preoperative BMD and postoperative periprosthetic BMD changes on the stability of either cemented or cementless cups. In contrast to the present study, Tabori-Jensen et al. did not observe a statistically significant difference in proximal cup migration for either fixation type. They therefore concluded that the results were inconsistent with our results. However, taking a closer look at the results of Tabori-Jensen et al., the intergroup difference in proximal migration for the preoperative BMD groups was at or over 0.1 mm for both cup types at individual timepoints. This difference is well within the effect-size confidence interval (95% CI: 0.06-0.37 mm) of the current study suggesting good congruence with our results, instead. Further, the overall mean magnitude of migration was considerably less in the proximal direction for Tabori-Jensen et al. which further explains the comparatively smaller difference between the groups suggesting that the statistically nonsignificant finding was merely a question of insufficient statistical power.

A critical difference between the study conducted by Tabori-Jensen et al. (2020) and the present study relates to the use of adjunct fixation. Our study protocol involved the insertion of pegs as fixation augments for all patients whereas the cementless Avantage DM cup used by Tabori-Jensen et al. relied on press-fit fixation alone. In an interesting parallel with the fourth study of this thesis, we could hypothesize that by constricting migration on some axes (e.g., rotation axes and X- or Z-translation) the pegs may have caused more of the 3D micromotion of migrating cups to align with the Y-axis. Conversely, it could be speculated that press-fit fixation alone allows for a greater variation in the direction of significant micromotion between individuals. Consequently, more of this significant micromotion may be leaked to other degrees of migration while reducing the observed difference aligning with the Y-axis. Indeed, Tabori-Jensen et al. noted that in the low preoperative BMD group, statistically significant migration differences could be observed between the fixation regiments on X-translation as well as Z- and Y-rotation. The cementless group, as contrasted by the preoperative BMD allocation, was not analyzed for these differences by Tabori-Jensen et al. Still, a similar difference can be speculated based on the provided graphs which suggests that there indeed is a 3D difference in migration between the groups. Of course, this line of discussion is highly speculative but, overall, it seems that a multivariate analysis might have revealed interesting phenomena from the data of Tabori-Jensen et al.

The methodology for determining the preoperative BMD was a crucial consideration in the present study. The indication for the index operation in our patient population was advanced osteoarthritis which is the most common indication for total hip arthroplasty (Ferguson et al. 2018). Some evidence points to a negative

correlation between osteoarthritis and osteoporosis though the matter is still debated (Shen et al. 2013, Im and Kim 2014). In apparent contrast with this notion, high incidences of osteoporosis have been reported for female patients scheduled for THA due to osteoarthritis (Glowacki et al. 2003, Mäkinen et al. 2007). Also, the subchondral sclerosis related to osteoarthritis may mask the diagnosis of osteopenia or osteoporosis if only local BMD near the affected joint is used (Lingard et al. 2009). Therefore, we did not use local BMD in the present study but rather used the lowest BMD value observed at several sites distant to the affected joint.

A related consideration is whether the increased migration of the acetabular cups was directly related to the decreased systemic BMD or, rather, the post-operative changes in periprosthetic BMD. This issue is relevant because patients with osteoporosis may be prone to greater decreases in periprosthetic BMD in the early post-operative period (Alm et al. 2009). Also, preventing this type of BMD loss is possible with antiresorptive drugs which paints it as an enticing target for a pharmacological intervention (Wilkinson et al. 2001). Unfortunately, the data from the present study does not shed light on the importance of post-operative BMD changes for the early stability of the cup. The fact that zoledronic acid, administered postoperatively, seems to reduce cementless acetabular cup migration underlines the importance of postoperative periprosthetic bone loss (Friedl et al. 2009). After the conduction of the present study, Tabori-Jensen et al. (2020) also examined the correlation between early cementless cup stability and post-operative periprosthetic BMD changes. They reported trending but statistically nonsignificant results for cementless cups with a group size of 30. Unfortunately, their study did not examine the interrelation between preoperative BMD and the postoperative changes in periprosthetic BMD.

A principal strength of the current study was the highly screened, homogenous patient population consisting of consecutive patients. Conditions and medications that could have influenced bone metabolism were perceived as potential confounders and used as exclusion criteria. Another major strength of this study was the use of the manufacturer-modified acetabular cup which facilitated marker-based RSA. No less than 13 RSA markers were attached to each acetabular cup, including the peg markers. The failure of marker-based RSA due to occlusion of both reference and implant markers is a recognized issue – especially in the case of metal-backed acetabular cups (Baad-Hansen et al. 2007). Model-based methods may help in case of occluded markers but even then, as most acetabular cups have an axis of almost perfect symmetry, determination of rotation may be infeasible on all axes (Jacobsen et al. 2018). In the current study, visualizing the markers placed at the cup opening proved often challenging in the clinical setting, as expected. Still, utilizing the markers attached to the pegs, marker-based RSA of both translations and rotations was successful in 34 out of 40 patients. This number of dropouts, while definitely

leaving room for improvement, compares well to similar marker-based RSA studies (Kärrholm and Snorrason 1992, Carlsson et al. 2006, Wolf et al. 2010, Naudie et al. 2013, Lazarinis et al. 2014, Jacobsen et al. 2018).

The main limitation of our study was the small number of patients with normal BMD. Due to the surprisingly high prevalence of osteopenia and osteoporosis among the study participants only twelve patients with normal BMD were recruited which is below the recommended group size for an RSA study (Valstar et al. 2005). Still, the results of the normal BMD group, suggesting low risk of revision due to aseptic loosening, were well in line with the long-term registry results of the ABG II cup in patients with a similar age range (Mäkelä et al. 2010). Additionally, the normal BMD group showed no outliers and had small variances for all migration parameters which allowed accurate estimation of migration parameter means despite the small sample size. Another limitation related to the recruitment strategy was imposed by ethical considerations. All patients with severe osteoporosis (T-score < -3.5) were treated with antiresorptive medication and therefore excluded from the present study. Consequently, the bone quality of patients analyzed with RSA was not entirely representative of the sampling population which likely had somewhat worse mean BMD.

We must also recognize a relevant limitation that relates to use of RSA-measured stability for predicting the clinical outcome of patient groups stratified according to systemic BMD. The body of evidence tying early RSA-measured migration and later aseptic loosening together is considerable, as discussed previously. Still, the relevant meta-analyses were made as a function of different implant designs and not patient related determinants for the risk of revision (Pijls et al. 2012a). This places some uncertainty as to whether the observed increased migration in our study for the low BMD group correlates with the risk of revision as expected. To the author's knowledge the performance of RSA as a predictor for the long-term clinical outcome as a function of various patient-related factors has not been studied systematically. We must therefore conclude that it is possible the osseointegration in the low BMD group was only delayed with little or no clinical impact.

Finally, our observational study setting, despite strict inclusion and exclusion criteria, was susceptible to confounders (Grimes and Schulz 2002). Compared to randomized interventional studies, stratifying patients according to patient-related factors is inherently prone to bias. For example, it has been previously noted that low BMD correlates with a lower physical activity and age (Carter and Hinton 2014). On the other hand, lower physical activity, also associated with aging, seems to reduce the risk of acetabular cup revision (Flugsrud et al. 2007). Thus, the effect of low BMD could have been masked in the present study due to the age difference between the groups that was a likely result of the patient stratification strategy. For this reason, the obvious confounders like age and BMI were controlled for in the data

analysis. However, it is still possible that some unrecognized or unmeasured sources of bias existed in our study population.

7.2 Study II: MBRSA of the Accolade II Stem

When planning an RSA-study, confirming the high level of accuracy and clinical precision of the method is essential. The second study of the thesis examined the accuracy and precision of model-based RSA in the migration analysis of a parallel-sided tapered-wedge femoral stem using a combined head-stem model. The *in vitro* accuracy of the method was in line with the marker-based method and the clinical precision was comparable to precision values reported in the recent literature (**Table 1**). Under the assumption of no significant bias, the achieved good clinical precision also indirectly suggests good accuracy in the clinical setting (Ranstam et al. 2000).

Besides the typical error sources in conventional RSA, the model-based method is also susceptible to errors caused by the specific implant design and inaccuracies of the implant surface model (Seehaus et al. 2013, 2016). For optimal results with the model-based method, an implant specific reverse-engineered surface model would be preferred. This is, however, often implausible for logistical and financial reasons (Seehaus et al. 2016). In the present study, the achieved good accuracy and precision suggest that the quality of the manufacturer-supplied CAD models was satisfactory for this implant type. Further, the specific shape of the Accolade II stem also seems suitable for analysis with model-based RSA.

The main limitation of the study protocol was that the accuracy of the Z- and X-rotation axes was not evaluated due to the associated technical difficulty. However, considering that these rotations occur along the long axis of the implant, it should be expected that their accuracy is at least as good and likely much better than the accuracy of the Y-axis rotation – even when using the combined head-stem model (Prins et al. 2008). Also, we did not use either a soft-tissue phantom or an anatomically shaped femur phantom due to technical concerns of possible experimental errors. This might affect the comparability of our accuracy results with dissimilar phantom setups.

The observed confidence interval for clinical precision was more than three times larger for the Z-axis translation compared to X- or Y-axis translation. This is well within the difference that would be expected for the precision of in-plane (X- and Y-axis translation) and out-of-plane (Z-axis) translations using model-based RSA (Seehaus et al. 2013). This effect is related to the use of a uniplanar cage and less than 90 degrees convergence of the x-ray beams (see 2.2.1.2). Thus, motion on the Z-axis causes relatively smaller changes in the radiographic projection causing pose-estimation on this axis to be less precise. In a similar fashion to Z-translation, the out-of-plane rotations on X- and Y-axis were expected to have lesser rotational

accuracy and precision. The simultaneous presence of relative rotational symmetry of hip implants along their long axis leads to the Y-axis rotation usually having even worse precision than X-axis rotation. In the present study the use of a combined head-stem model may have aided in maintaining the relatively good precision of Y-axis rotation (Prins et al. 2008).

An interesting phenomenon associated with model-based RSA is that the precision of the method may be affected by large rotations and translations of the implant relative to the radiographic setup (Kaptein et al. 2003, Lindgren et al. 2020). This phenomenon is accentuated with the use of implant surface models of suboptimal quality (Kaptein et al. 2003). The fact that marker-based RSA is affected less by this phenomenon (Lindgren et al. 2020), confirms the precision of the established methodology for estimating the configuration of the radiographic setup. Therefore, we can conclude that this error is related to the model-based methodology: With near identical positioning the same parts of the surface model are used to generate the digitally reconstructed radiographs necessary for the model-based method and, thus, any potential errors due to inaccuracies of the surface model are zeroed out in successive comparisons (Kaptein et al. 2003). On the other hand, with extreme translations and rotations different parts of the surface models are used to generate the digitally reconstructed radiographs and the errors in the surface model are transferred to the migration estimates. In the present study the relatively good accuracy of the Y-rotation axis, despite quite large rotations (10 degrees), further supports the good accuracy of the manufacturer-supplied surface models. Still, for future phantom studies using model-based RSA, controlled rotational and translational displacement of the entire phantom setup, within the likely clinical range, should be a consideration.

7.3 Study III: DLRSA of Internally Fixated Femoral Neck Fractures

The aim of internal fixation is primary fracture healing (primary cortical union) that is only enabled by stable fracture fixation (Marsh and Li 1999, Marsell and Einhorn 2011). Consequently, characterizing and optimizing the early stability of femoral neck fractures has been a focus of considerable research efforts in preclinical biomechanical studies (Cha et al. 2019). These preclinical efforts are at a stark contrast with the paucity of related clinical research. Only few authors have even attempted to characterize the stability of femoral neck fracture fixation *in vivo* due to the evident lack of viable research methods (Elmerson et al. 1987). DLRSA has shown promise in the clinical study of fracture healing but due to the related technical difficulty it's validation for each fracture type separately has been recommended (Chehade et al. 2009). The third sub-study was specifically designed to validate

DLRSA as a clinical research tool for femoral neck fracture stability after internal fixation.

Inducible micromotion at the fracture site could be detected with the deployed DLRSA methodology. Underlining the potential clinical relevance of DLRSA, a surprisingly strong association between the six-week inducible micromotion and later clinical non-union was observed. Lack of rotational stability at the fracture site seemed especially detrimental for the clinical outcome. Each of the five failure cases predicted by DLRSA were outliers in rotational compliance either at baseline or at six weeks. This finding closely corroborates the notion that rotational movements and the related shear at the fracture site are potent inhibitors of fracture healing (Gaston and Simpson 2007).

The methodology for generating and measuring loading directed at the fracture site is critical for any DLRSA study (Chehade et al. 2009). Confirming the relevancy of the used methodology for fracture-site loading, our DLRSA results were closely in line with related cadaver studies. The axial compliance in cadaver models of femoral neck fractures fixed with cannulated screws has been observed to vary between 1 to 10 μm / N (Cha et al. 2019). Assuming that healed fractures should have strength equaling that of healthy bone, we would have expected to see compliances of approximately 0.5 to 2 μm / N for the fully healed fractures (Sjöstedt et al. 1994, Miura et al. 2017). These limits corresponded with the compliance we observed with DLRSA for all fractures at baseline and for the healed factures at 52 weeks (**Figure 10**).

Interestingly, conventional RSA showed gradual sustained migration until the end of the 52-week follow-up even for fractures that achieved union (**Figure 11**). In previous RSA studies a similar observation concerning both screw and hook-pin fixation has been reported: gradual migration in healing fractures could be detected for up to nine to 12 months after surgery (Ragnarsson and Kärrholm 1991, Ragnarsson et al. 1993). Meanwhile, the DLRSA results seem to suggest that for most healing fractures near maximal stiffness was achieved by 12 weeks, already. The relationship between fracture union and fracture-site stiffness has not yet been determined for femoral neck fractures. However, in an ovine model of externally fixated tibial fractures, this initial phase of rapidly increasing stiffness was associated with achieving union (Claes and Cunningham 2009). If this finding applies to femoral neck fractures as well, we could hypothesize that a considerable portion of the later occurring sustained micromotion is a result of fracture-site remodeling. Therefore, sustained micromotion may provide only a lagging indicator of time to fracture union.

Due to the main goal of validating DLRSA, the timing of baseline conventional RSA (within three days postoperatively) was not a critical consideration for the current study. It has been previously shown, that significant migration of the femoral

head occurs already during the first post-operative day even before weight bearing is allowed (Ragnarsson et al. 1993). Consequently, the observed sustained RSA migration in our clinical cohort was much lower (translation vector length of 4.26 mm) than that detected in comparable RSA studies where baseline imaging was performed on the day of the surgery (Ragnarsson and Kärrholm 1992, Ragnarsson et al. 1993, Mattsson and Larsson 2003). Indeed, the results based on the conventional radiography measurements were much more in line with the expected scale of migration (mean FNS of 9.2mm). This magnitude of migration is also in line with the reports of other research groups that used standard AP radiographs for migration analysis (Zlowodzki et al. 2008, Zielinski et al. 2013). As to the clinical outcome, the revision rate of 31% in our relatively small patient population was overall in line with the results reported in a multi-center trial concerning femoral neck fractures (Nauth et al. 2017).

As discussed previously, DLRSA may be precise enough so that flexibility at a healing fracture site and surrounding bone may play a significant role in the amount of migration that is detected. This may result in significant inducible migration being detected especially at the group level even when the fractures are healing as expected. Therefore, comparing the inducible micromotion to an estimated cut-off point has been recommended (Wilson et al. 2010). For femoral neck fractures such limits are unknown. The current study was not powered or, indeed, designed to estimate a clinically relevant cut-off point or timing between acceptable and non-acceptable inducible micromotion. Thus, the precision limits of the method were used as cut-off points. While these limits seemed highly relevant in our small patient cohort, using the precision limits is likely not the optimal approach which should be a consideration for future research. Similarly, the optimal timing of the DLRSA examinations is still an open question but our results seem to suggest that for most cases the clinical course to union or non-union has already been determined by six to twelve weeks post operatively.

The main limitation and obstacle faced by the current study was the difficulty in reliably placing a sufficient number of RSA markers into the femoral head. The methodology for RSA of femoral neck fractures has been described previously (Ragnarsson and Kärrholm 1991). Analogous to the previously described technique, three markers could be implanted into the femoral head through the cannulated screws in the present study. However, implanting markers with satisfactory spread and visibility proved much more difficult with the sliding hip screw. This led to early cessation of the planned randomization. As a related limitation, we accepted larger than ideal condition numbers. This reflects the difficulty of achieving an optimal scatter of RSA markers given the small spatial dimensions of the femoral head and the difficulty in the application of the markers. As recommended for studies where ideal condition numbers cannot be achieved, we confirmed the precision of the

measurements for each patient using double-measurements (Valstar et al. 2005). Improved instrumentation and, if necessary, cadaver experimentation are called for to ensure safe implantation of markers in future femoral neck fracture studies where predrilled holes or fixation hardware don't allow for trivial marker placement.

As a further limitation, the small sample size and the explorative nature of the study restrict the external validity of the results. The finding that inducible micromotion could be tied to the clinical outcome already at such a small sample size was a positive surprise. Still, due to the small sample, the effect size (i.e., risk or odds ratio of non-union) of the observed phenomenon has considerable uncertainty. Finally, it must be stressed that, despite the strong association between DLRSA results and non-union, the value of DLRSA does not lie in *predicting* the clinical outcome of femoral neck fracture healing. This is in contrast with the role of RSA in the study of either hip or knee implants. Even if the predictions made by DLRSA were faultless as to union or nonunion, the same information of the clinical outcome can be gained from the clinical cohort with just one to two years' clinical follow-up. The true utility of DLRSA for the study of femoral neck fractures lies instead in better understanding the biomechanics of the fracture site – perhaps allowing the differentiation of *why* some individual cases proceed to union and some to non-union.

7.4 Study IV: Multivariate LMM Analysis of Accolade II stem

RSA produces migration data with six degrees of freedom when using rigid-body kinematics. A typical RSA study also has multiple follow-ups. Still, the common practice among RSA studies seems to be to analyze the data using only univariate methods. This may be related to the perceived complexity of multivariate data analysis and, also, the relative novelty of the relevant multivariate analytical methods. The fourth sub study of this thesis presented a multivariate analysis of clinical RSA data using LMM. The study was designed as only a practical example of the statistical method but incidentally also provided clinically interesting insights of the utilized RSA data.

When applied to the adopted clinical data, multivariate LMM revealed a trending difference in the migration patterns between the two groups. The inspection of the LMM effect-slices demonstrated how the results of the multivariate model could be broken into univariate terms: the intervention group showed a greater stem Z-rotation. This difference seemed to develop during the first 12 weeks. No further time-related changes were evident between the groups (**Figure 12**). These observations suggest that denosumab does, indeed, modify the early implant stability

of the cementless femoral stems – if only during the immediate postoperative period up until three months.

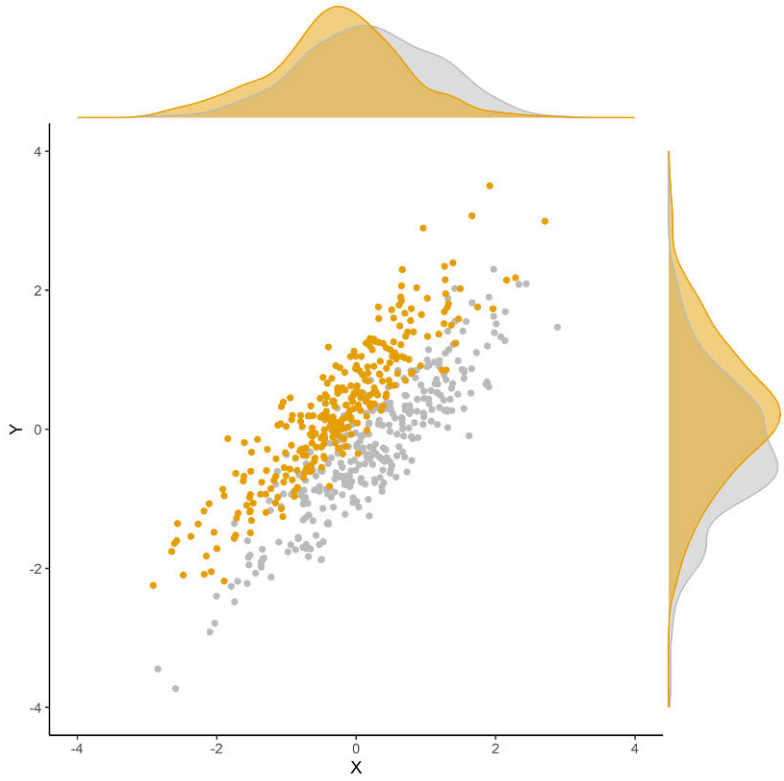


Figure 14. A scatterplot with marginal distributions of a simulated bivariate-normal response with strong autocorrelation for two distinct groups. The difference between the groups is very small in univariate terms, as demonstrated by the marginal distributions, although in bivariate terms the difference is nearly categorical.

A Monte Carlo simulation was used to demonstrate that using multivariate analytical methods for RSA data analysis does not necessarily result in reduced statistical power. When the mean difference in migration aligns perfectly with the axis of interest, intuition would dictate that a univariate analysis on this axis would have superior statistical power. However, as demonstrated by the Monte Carlo simulation, the opposite may in fact be true. The critical factor is to consider that, even if the intergroup difference manifests only on a single degree of freedom, the confidence intervals of the groups are also multivariate. Thus, a univariate analysis may underestimate the significance of the intergroup difference (**Figure 14**).

In data where there is zero covariance for the response's degrees of freedom, a multivariate analysis is, of course, somewhat conservative compared to analyzing

only an individual degree of freedom. This is a necessary concession for maintaining the aspired alpha-level when comparing multiple degrees of freedom simultaneously. Even in this worst-case scenario for the multivariate LMM, the method maintained significantly higher statistical power (48.1%) compared to analyzing total translation and total rotation (24.7%). Still, the complete absence of covariance for the response variables is an entirely hypothetical notion that is exceedingly unlikely in clinical data. Thus, the statistical power of 48.1% can be considered a lower-bound for the method with likely much better statistical power being observed in practice if the core assumptions of the method hold.

A key advantage of the applied statistical method was protecting the statistical inferencing from the risks associated with statistical multiplicity. This point was concretely demonstrated by the empirical examination of the alpha-levels. However, the difference is even more pronounced if one considers the fact that a typical RSA study has multiple follow-ups. Indeed, in the utilized clinical data, 18 distinct RSA datapoints were available for each study subject. Using univariate methods, comparison of such a large number of datapoints between two groups would have been highly dubious – with or without attempts to correct the results for multiplicity. Dimensionality reduction techniques can be used to facilitate univariate data analysis. However, at least in the case of total translation and total rotation, the statistical power of these methods was alarmingly close to the false positive rate: 24.7% compared to 9.5%. This observation underlines the fact that, even when used in conjunction, total translation and total rotation provide poor measures for detecting significant migration differences.

High quality statistical analysis is only the first step of inferencing a clinical significance. On a fundamental level, statistical significance implies merely that the analytical result is most likely repeatable and therefore generalizable to the defined sampling population (Ranganathan et al. 2015). Delineating statistical and clinical significance requires further correlating the observed difference with clinically meaningful effects and effect sizes (Rosenberg et al. 2012). Arguably, the existing understanding of the RSA measured micromotion and its causalities is insufficient to decide on the optimal analytical model or method.

For LMM, a critical question is whether a *linear* statistical model is suited for detecting clinically relevant differences in RSA data. In other words, the question is related to whether there are valid grounds to assume a linear relationship between group mean migration and the risk of revision. Aspenberg et al. (2008) have shown that their population of cemented acetabular cups was divided to stable non-migrators with a mean migration close to zero and loose migrators with non-zero migration. They also hypothesized that the loose migrators were likely, as a group, at a higher risk of revision. This hypothesis would seem to suggest that the risk of revision has somewhat of a dichotomous distribution as a function of early

micromotion. Indeed, Aspenberg et al. proposed that the proportion of stable and unstable implants in a population could be a superior measure of the clinical outcome.

It could be argued that assuming a rigid dichotomy is an oversimplification of the complex processes and causal interrelations that lead to RSA measured migration and revision due to aseptic loosening. It is likely that on individual level there is an overall gradient for RSA measured migration and the risk of revision in addition to a possible dichotomy. In other words, while risk of revision is perhaps greatly increased at or around some magnitude of migration or change in migration, even below or beyond such limit greater migration infers relatively greater risk of aseptic loosening and *vice versa*. Further, when dichotomous data is averaged in a study sample, the observed relationship between *mean* migration and risk of revision due to aseptic loosening may be surprisingly linear. Statistically speaking, if the sampling population is divided into two distributions with different means, the observed sample mean would correlate linearly with the proportion of observations from either distribution.

Further, considering this issue in a broader context, many if not most natural phenomena are inherently nonlinear. Yet, analytical methods based on linearity have been highly successful in modeling data even with known non-linear relationships (Yu 2010). When nonlinear models are utilized the choice of the specific model should ideally be based on observed patterns in the data or a theoretically justifiable basis. In their absence a linear model is often justifiable in the interest of simplicity (Yu 2010). This seems to further support the relevance of a linear analytical model for RSA measured migration – at least until a better understanding of the phenomena underlying RSA measured migration is gained.

The main limitation of the present study was the fact that it performed a retrospective reanalysis on previously published clinical RSA data. Further, the clinical data only included postmenopausal females scheduled for THA due to advanced osteoarthritis. These considerations reduce the generalizability of our findings. Further, the 3D difference in migration between the groups was only trending but we nevertheless proceeded to the inspection of the model effect-slices.

Another critical limitation of the present study has to do with the role of outliers in RSA data. The main body of evidence on the clinical significance of RSA measured migration is focused on population mean migration, as discussed previously. Outliers can affect the group mean estimates and results of parametric statistical methods considerably. The question then is raised if outliers in RSA data are enabling or, rather, impeding us from arriving at clinically relevant conclusions about RSA data. If outliers in RSA data are mostly a product of the dichotomy discussed above, they should not be excluded from the data as this would distort the relationship between loose implants and increased migration. If, however, the

outliers are a product of technical errors or excessive biological variability, their exclusion may be justified to arrive at a more accurate estimate of the group mean. Currently the evidence is insufficient to decide on the significance of RSA outliers for the data analysis and the debate is ongoing (de Vries et al. 2014).

In the case of the present study, the clinical interest in the adopted RSA data was focused on the very early implant stability. The likelihood of outliers due to migration data dichotomy – developing only later due to loose migrators (Aspenberg et al. 2008) – was therefore judged to be insignificant. We consequently excluded outliers using a previously published definition in the same patient cohort (Nazari-Farsani et al. 2020, 2021). As demonstrated previously, these patients seemed to constitute a clinically distinct sub-group further justifying their exclusion (Nazari-Farsani et al. 2021). Arguably, this allowed for a more accurate assessment of early implant stability in the immediate postoperative period. Nevertheless, the exclusion of outliers may have introduced bias in the data.

To minimize concerns of selection bias, we wished to adhere to the previously published definition of outliers in our data. The definition published by Nazari-Farsani et al. (2020) used Y-axis translation and rotation to identify outliers. To facilitate the multivariate analysis, focused not only on the Y-axis, one further patient was excluded due to nearly an order of magnitude greater X-rotation compared to that of any other patient. Still, this definition of outliers was hardly ideal from the viewpoint of multivariate LMM. Future studies using multivariate LMM, should ideally use the proper model diagnostic tools for identification of multivariate outliers in the data (SAS Inst. Inc 2009).

8 Future Prospects

Concerning the first study, the logical follow-up question is if the implied high risk of acetabular cup revision due to aseptic loosening for female patients with low systemic BMD holds true at closer inspection. Using registry data, this hypothesis could perhaps be tested by systematically identifying subpopulations with comorbidities or demographics that imply a low systemic BMD. If inspection of such subpopulations were to reveal a higher-than-expected revision rate due to aseptic loosening, when compared to matched controls, it would serve as confirmation of our RSA-based prediction. Such comparisons are not without difficulty, however, due to the various confounding factors associated with BMD, as discussed previously.

The existing evidence on patient-related determinants for RSA-measured implant stability is few and far between. This is somewhat curious considering the vast potential benefits associated with a better understanding of the RSA-measured migration and its causes: First, one of the principal strengths of RSA is the high accuracy and precision of the method allowing for small sample sizes. However, this advantage is lost if the measurements are overwhelmed with biological – i.e., patient-related – variation. Even a study focusing solely on implant-related determinants of early stability would benefit from controlling for the patient-related factors. The advantages come in the form of lesser unexplained variation in the data and, consequently, more accurate estimates for the main effect. Second, the patient-related factors for early implant stability are not set in stone and therefore represent interesting potential targets for intervention. In any case, a systematic study of patient-related determinants for early implant stability would likely have an important impact on future RSA studies of implant micromotion.

Regarding potential interventions preceding THA with a cementless acetabular cup, the prospect of antiresorptive treatment seems all the more interesting considering the findings of the first sub study. The impact of antiresorptives could be first studied with RSA. In fact, already before the conduction of our study, an intervention with zoledronic acid had shown promise for optimizing the stability of a cementless acetabular cup (Friedl et al. 2009). An important consideration also relates to the optimal timing of the antiresorptive treatment for the prevention of

RSA-measured migration: if preoperative BMD is an independent predictor of early implant migration and not just a surrogate measure for later periprosthetic bone loss, starting the antiresorptive treatment well in advance of the index operation might maximize the potential benefits – especially among patients with low preoperative BMD. Thus, carefully characterizing the impact of both pre- and postoperative BMD in relation to RSA-measured implant stability remains an interesting research topic for optimizing the timing of the intervention.

The second sub study successfully validated the model-based method for a cementless tapered-wedge femoral stem. This study subsequently pawed way for a clinical trial using the described methodology (Aro et al. 2019). These studies represented the first time that RSA based on implant surface models was utilized in our laboratory. Going forward this work provides an important steppingstone for moving the laboratory over to the model-based methodology which will facilitate future RSA-research. Still, the development of the RSA methodology is ongoing and progressing at a rapid pace. Thus, any future clinical RSA studies in our laboratory will require conducting a similar phantom study for the validation of the method.

The obvious use-case for DLRSA in the study of femoral neck fractures is to confirm that the accumulated *in vitro* research data translates to *in vivo* as expected for various internal fixation techniques. Additionally, carefully characterizing the biomechanical properties in the immediate postoperative period might provide further insights as to the risk factors of non-union. This might facilitate the development of improved internal fixation techniques and potentially help better choose the femoral neck fracture patients that will benefit from internal fixation over other treatment modalities. From a wider perspective, DLRSA may yet prove a key research method for various fracture types and even a gold standard for defining fracture union in the clinical setting.

Beyond validating DLRSA as an *in vivo* research tool for femoral neck fractures, the second sub study also pointed to an interesting future research avenue with regards to the rotational stiffness at the fracture-site. The biomechanics of femoral neck fracture internal fixation have been studied in numerous *in vitro* studies (Cha et al. 2019). However, these studies have typically focused on linear displacement of the femoral head in response to an external load. The findings of the second sub study seem to point to the fact that rotational stability may also be an important consideration – especially for internal fixation with cannulated screws. Therefore, optimizing the fixation for rotational stability may represent an interesting research prospect. DLRSA is at present the only tool with the ability to accurately quantify the rotational stability of the fracture *in vivo.* Even *in vitro* DLRSA could perhaps be used to facilitate robust and precise measurements of rotation in cadaver and saw-bone fracture models.

The potential benefits of multivariate RSA data analysis were well demonstrated by the fourth sub study. Going forward it is important to recognize that these benefits are not limited to improved statistical power and robustness of statistical inferencing. The ability to identify and differentiate multivariate migration patterns in RSA data enables the examination of much more varied and complex hypotheses regarding the significance of migration. Using multivariate statistical methods these hypotheses can trivially encompass time-related and measurement axis-related changes and correlations. As a prime example, the time-related changes in migration may be valuable modulators of the clinical outcome, as discussed previously.

An even more interesting question based on a multivariate interpretation of RSA data is whether identifying correlations between measurement axes could have value in the data analysis. For example, a relatively high degree of subsidence may be acceptable for cementless femoral stems (de Vries et al. 2014, van der Voort et al. 2015). However, whether the subsidence occurs in conjunction with, or absence of, stem valgus rotation may provide important insights as to how much periprosthetic resorption is occurring near the lesser trochanter. A greater stem valgus rotation, as observed in the analysis of the clinical trial data, likely implies a lesser degree of bone resorption thus pushing the implant into valgus as a result of the subsidence. Therefore, the presence or absence of a correlation between the stem subsidence and valgus rotation could hypothetically be a surrogate measure for later risk of revision. Whether these types of multivariate phenomena truly exist in RSA data remains to be seen but without multivariate analytical methods and thinking they are unlikely to be discovered in the first place.

Perhaps the most interesting research prospect, as a direct continuation of this thesis, is formulating a mathematical model of RSA-measured micromotion based on the current understanding of the causalities that lead to implant migration. Having the ability to simulate RSA data under various assumptions would have obvious benefits: For example, if dichotomous RSA data could be simulated based on the work of Aspenberg et al. (2008) this would allow comparing various statistics and analytical models empirically. Using such a method we could identify the best statistical model or measure for predicting the long-term revision rates of an implant – under the assumption of dichotomy. Further, a comprehensive mathematical formulation of implant migration could be used to simulate various experimental designs of RSA studies in advance. This would allow the optimization of the experimental design in such a way that a maximum of information is gained using a minimum of research resources. Of course, such a mathematical model would be imperfect due to the gaps in our present understanding of implant migration. However, even imperfect models of the real world are often an important step in forming testable hypotheses for future research.

Overall, RSA remains a resource intensive research method that is at present unsuitable for routine clinical follow-up of individual patients. In the future, the development of marker-free RSA methodology and CT-based alternatives could yet prove useful tools even outside dedicated research. However, at present there is a scarcity of data on the impact of RSA-measured migration at the individual level and whether the additional information provided by RSA cloud benefit day-to-day clinical decision-making. Still, as demonstrated by this thesis, even when constrained to clinical research the potential of RSA, as the enabler of innovation in the development of orthopaedic implants and related interventions, is tremendous.

9 Conclusions

Based on the studies constituting this thesis the following conclusions can be drawn:

1. As hypothesized, preoperative systemic bone mineral density represents one of the underappreciated patient-related factors dictating the early stability of cementless acetabular cups in the female population. Crucially, this patient-related difference in early implant stability can be detected using the deployed marker-based RSA methodology.
2. Model-based RSA is sufficiently accurate and precise for measuring migration of the tapered wedge Accolade II cementless femoral stem. The precision is also maintained *in vivo* indicating that the method is comparable to marker-based RSA and that the method is valid for future clinical research of the Accolade II stem.
3. With the presented methodology, DLRSA can be utilized for the clinical study of femoral neck fractures fixed with cannulated screws. Increased fracture-site compliance and continuation of inducible micromotion beyond baseline is an early indicator resulting nonunion. DLRSA is a unique research tool for assessing femoral neck fracture fixation stability *in vivo* and could be applied to comparison of different fixation methods.
4. As a proof-of-concept, the application of multivariate LMM for the analysis of RSA data revealed an unexpected, complex 3D change in the migration patterns of a cementless femoral stem in response to the antiresorptive medication, denosumab. Future RSA research could benefit from the application of multivariate LMM for exhaustive analysis of RSA data. The multivariate LMM can provide greater sensitivity for detecting complex 3D migration patterns while also minimizing the risk of false-positive analysis results.

Acknowledgements

This thesis and the related studies were conducted in the Orthopaedic Research Unit of the Department of Orthopaedic Surgery and Traumatology in Turku University Hospital and the University of Turku. I wish to thank Professor Emeritus Hannu Aro for allowing me to utilize his facilities at the University of Turku for carrying out the work and research related to this thesis.

I'd like to express my gratitude to my supervisors Professor Emeritus Hannu Aro and Adjunct Professor Niko Moritz. Hannu, your enthusiastic and stoic support was crucial – more than once – in bringing this thesis all the way to its conclusion. It has been a privilege to work under your supervision. Niko, you have taught me invaluable lessons on subjects ranging all the way from the nitty-gritty details of data gathering on to how to conduct oneself as an objective scientist. I am deeply indebted to you both.

I'm also grateful for the in-depth advice and feedback from the pre-examiners of this thesis, Professor Miika Nieminen and Adjunct Professor Teemu Moilanen. Your feedback on both the most minute details and also the coherence of the thesis as a whole were a testament to the time and effort you put in to thoroughly evaluate my thesis. Your critical insights were invaluable for finalizing this thesis.

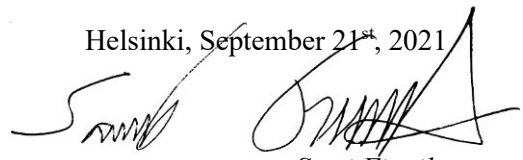
This thesis would not have been possible without the help and support of my coauthors: Dr Sanaz Nazari-Farsani, Dr Jessica Alm, Docent Kimmo Mattila, Docent Erkki Svedström, Dr Niko Strandberg, and Eliisa Löyttyniemi. I'm grateful to you all for your hard work, support, and advice during each of the sub studies as well as during the writing of this thesis. Additionally, I wish to thank Dr Petteri Lankinen and Clinical Study Coordinator Satu Timlin: although we worked together only very briefly your support and encouragement were invaluable at a time when I still had very little idea of what I was getting into.

This thesis was made possible by personal grants from Suomen Artroplastiayhdistys (SAPLY), the Doctoral Programme in Clinical Research (DPCR) at the University of Turku, and the TULES-domain of the Turku University Hospital. I also wish to acknowledge all the help and support I've received from the various members of the DPCR at the University of Turku.

I also wish to extend my heartfelt gratitude to my mom and dad: Riikka and Risto. I think it's fair to say that, growing up, I rarely stuck with the beaten path. Regardless, you've always managed to provide me with your unwavering support – even during this project. Dad, you've always amazed me with your uncanny ability to sift the essential from the nonessential. To the best of my abilities, I've tried to achieve the same in both science and life in general. Mom – my loving, harebrained mother – you have taught me what it means to be creative: it's something that cannot be forced and must be given its time – even if it sometimes means dropping everything else on the schedule. In my humble experience the very best of scientific insights require a good deal of creativity.

Finally, Tiina: Quite a few PhD students before me, and in all likelihood many more PhD students will after me, have lamented over the toll that their thesis project has exacted on their families. I wish I wouldn't have to do the same but even as I write this you're taking on more than your fair share of chores and tasks around the house while looking after both Sonja and the baby. I've also long since lost count on how many text passages and emails I've asked you to proofread for this thesis. Yet, I have not once heard you complain. Quite the contrary, you have time and again encouraged me to continue my work. Without your loving support I could not possibly have completed this thesis. As the saying goes, actions speak louder than words and I'm humbled beyond words by the message your actions convey. No matter the challenges that lie ahead I'm grateful to have you by my side.

Helsinki, September 21st, 2021



Sami Finnilä

References

Acklin Y, Jenni R, Bereiter H, Thalmann C, Stoffel K. Prospective clinical and radiostereometric analysis of the Fitmore short-stem total hip arthroplasty. *Arch Orthop Trauma Surg*. Germany; 2016;136(2):277–84.

Ahl T, Dalén N, Holmberg S, Selvik G. Early weight bearing of malleolar fractures. *Acta Orthop Scand*. England; 1986;57(6):526–9.

Ahl T, Dalén N, Selvik G. Ankle fractures. A clinical and roentgenographic stereophotogrammetric study. *Clin Orthop Relat Res*. United States; 1989;(245):246–55.

Alm JJ, Mäkinen TJ, Lankinen P, Moritz N, Vahlberg T, Aro HT. Female patients with low systemic BMD are prone to bone loss in Gruen zone 7 after cementless total hip arthroplasty: A 2-year DXA follow-up of 39 patients. *Acta Orthop*. 2009;80(5):531–7.

Armstrong RA. Recommendations for analysis of repeated-measures designs: testing and correcting for sphericity and use of manova and mixed model analysis. *Ophthalmic Physiol Opt*. Blackwell Publishing Ltd; 2017;37(5):585–93.

Aro HT, Alm JJ, Moritz N, Mäkinen TJ, Lankinen P. Low BMD affects initial stability and delays stem osseointegration in cementless total hip arthroplasty in women: a 2-year RSA study of 39 patients. *Acta Orthop*. Taylor & Francis; 2012;83(2):107–14.

Aro HT, Engelke K, Mattila K, Löyttyniemi E. Volumetric Bone Mineral Density in Cementless Total Hip Arthroplasty in Postmenopausal Women. *J Bone Jt Surg*. 2020;103(12):1072–82

Aro HT, Nazari-Farsani S, Vuopio M, Löyttyniemi E, Mattila K. Effect of Denosumab on Femoral Periprosthetic BMD and Early Femoral Stem Subsidence in Postmenopausal Women Undergoing Cementless Total Hip Arthroplasty. *JBMR Plus*. Wiley; 2019;3(10):e10217.

Aronson A, Hoist L, Selvik G. An Instrument for Insertion of Radiopaque Bone Markers. *Radiology*. The Radiological Society of North America; 1974;113(3):733–4.

Asar Ö, Bolin D, Diggle PJ, Wallin J. Linear mixed effects models for non-Gaussian continuous repeated measurement data. *J R Stat Soc Ser C (Applied Stat*. Blackwell Publishing Ltd; 2020;69(5):1015–65.

Aspenberg P, Wagner P, Nilsson KG, Ranstam J. Fixed or loose? Dichotomy in RSA data for cemented cups. *Acta Orthop*. 2008;79(4):467–73.

Augat P, Faschingbauer M, Seide K, Tobita K, Callary SA, Solomon LB, Holstein JH. Biomechanical methods for the assessment of fracture repair. *Injury*. Elsevier Ltd; 2014;45(SUPPL. 2).

Ausiello G, Marchetti-Spaccamela A, Crescenzi P, Gambosi G, Protasi M, Kann V. The Complexity of Optimization Problems. *Complex Approx*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1999. p. 1–37.

Ayers D, Greene M, Snyder B, Aubin M, Drew J, Bragdon C. Radiostereometric analysis study of tantalum compared with titanium acetabular cups and highly cross-linked compared with conventional liners in young patients undergoing total hip replacement. *J Bone Jt Surg Am Vol*. 2015;97(8):627–34.

Baad-Hansen T, Kold S, Kaptein BL, Søballe K. High-precision measurements of cementless acetabular components using model-based RSA: An experimental study. *Acta Orthop*. Taylor & Francis; 2007;78(4):463–9.

Baad-Hansen T, Kold S, Nielsen PT, Laursen M, Christensen P, Soballe K. Comparison of trabecular metal cups and titanium fiber-mesh cups in primary hip arthroplasty: a randomized RSA and bone mineral densitometry study of 50 hips. Acta Orthop. 2011. p. 155–60.

Baldursson H, Egund N, Hansson LI, Selvik G. Instability and wear of total hip prostheses determined with roentgen stereophotogrammetry. *Arch Orthop Trauma Surg*. Springer-Verlag; 1979;95(4):257–63.

Baldursson H, Hansson LI, Olsson TH, Selvik G. Migration of the acetabular socket after total hip replacement determined by roentgen stereophotogrammetry. *Acta Orthop Scand*. 1980;51(3):535–40.

Bandyopadhyay S, Ganguli B, Chatterjee A. A review of multivariate longitudinal data analysis. Stat. Methods Med. Res. SAGE PublicationsSage UK: London, England; 2011. p. 299–330.

Bergvinsson H, Sundberg M, Flivik G. Polyethylene Wear With Ceramic and Metal Femoral Heads at 5 Years: A Randomized Controlled Trial With Radiostereometric Analysis. *J Arthroplasty*. United States; 2020;35(12):3769–76.

Bhandari M, Chiavaras M, Ayeni O, Chakraverrty R, Parasu N, Choudur H, Bains S, Sprague S, Petrisor B, Assessment Group for Radiographic Evaluation and Evidence (AGREE) Study Group (AGREE Investigators Writing Committee). Assessment of Radiographic Fracture Healing in Patients With Operatively Treated Femoral Neck Fractures. *J Orthop Trauma*. 2013;27(9):e213–9.

Bhandari M, Guyatt GH, Swiontkowski MF, Tornetta P, Sprague S, Schemitsch EH. A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. *J Orthop Trauma*. 2002;16(8):562–6.

Bizzoca D, Vicenti G, Caiaffa V, Abate A, De Carolis O, Carrozzo M, Solarino G, Moretti B. Assessment of fracture healing in orthopaedic trauma. *Injury*. Elsevier Ltd; 2020;

Blom IF, Koster LA, Brinke B Ten, Mathijssen NMC. Effective radiation dose in radiostereometric analysis of the hip. *Acta Orthop*. Taylor and Francis Ltd.; 2020;91(4):1–395.

Bobyn J, Pilliar R, Cameron H, Weatherly G. The optimum pore size for the fixation of porous-surfaced metal implants by the ingrowth of bone. *Clin Orthop Relat Res*. 1980;NO. 150(150):263–70.

Bojan A, Jönsson A, Granhed H, Ekholm C, Kärrholm J. Trochanteric fracture-implant motion during healing - A radiostereometry (RSA) study. *Injury*. Netherlands; 2018;49(3):673–9.

Bontempi M. On the algebraic definition of total rotation in RSA. 2020;

Börlin N, Thien T, Kärrholm J. The precision of radiostereometric measurements. Manual vs. digital measurements. *J Biomech*. Elsevier; 2002;35(1):69–79.

Bragdon CR, Estok DM, Malchau H, Kärrholm J, Yuan X, Bourne R, Veldhoven J, Harris WH. Comparison of two digital radiostereometric analysis methods in the determination of femoral head penetration in a total hip replacement phantom. *J Orthop Res*. 2004;22(3):659–64.

Brodén C, Olivecrona H, Maguire GQ, Noz ME, Zeleznik MP, Sköldenberg O. Accuracy and Precision of Three-Dimensional Low Dose CT Compared to Standard RSA in Acetabular Cups: An Experimental Study. *Biomed Res Int*. Hindawi; 2016;2016:1–6.

Brodén C, Sandberg O, Sköldenberg O, Stigbrand H, Hänni M, Giles JW, Emery R, Lazarinis S, Nyström A, Olivecrona H. Low-dose CT-based implant motion analysis is a precise tool for early migration measurements of hip cups: a clinical study of 24 patients. *Acta Orthop*. Taylor and Francis Ltd; 2020;91(3):260–5.

de Bruin PW, Kaptein BL, Stoel BC, Reiber JHC, Rozing PM, Valstar ER. Image-based RSA: Roentgen stereophotogrammetric analysis based on 2D–3D image registration. *J Biomech*. Elsevier; 2008;41(1):155–64.

Budde S, Seehaus F, Schwarze M, Hurschler C, Floerkemeier T, Windhagen H, Noll Y, Ettinger M, Thorey F. Analysis of migration of the Nanos® short-stem hip implant within two years after surgery. *Int Orthop*. Germany; 2016;40(8):1607–14.

Cai R, Yuan X, Rorabeck C, Bourne RB, Holdsworth DW. Development of an RSA calibration system with improved accuracy and precision. *J Biomech*. Elsevier; 2008;41(4):907–11.

Callaghan JJ, Bracha P, Liu SS, Piyaworakhun S, Goetz DD, Johnston RC. Survivorship of a Charnley Total Hip Arthroplasty. *J Bone Jt Surg*. 2009;91(11):2617–21.

Carlsson L V, Albrektsson T, Albrektsson BEJ, Jacobsson CM, Macdonald W, Regnér L, Weidenhielm LR. Stepwise introduction of a bone-conserving osseointegrated hip arthroplasty using RSA and a randomized study: II. Clinical proof of concept—40 patients followed for 2 years. *Acta Orthop*. Taylor and Francis Ltd.; 2006;77(4):559–66.

Carter MI, Hinton PS. Physical activity and bone health. Mo. Med. Missouri State Medical Association; 2014. p. 59–64.

Cha YH, Yoo J Il, Hwang SY, Kim KJ, Kim HY, Choy WS, Hwang SC. Biomechanical evaluation of internal fixation of pauwels type III femoral neck fractures: A systematic review of various fixation methods. *CiOS Clin Orthop Surg*. Korean Orthopaedic Association; 2019;11(1):1–14.

Charnley J. ARTHROPLASTY OF THE HIP: A New Operation. *Lancet*. Elsevier; 1961;277(7187):1129–32.

Charnley J. Total hip replacement by low-friction arthroplasty. *Curr Orthop Pract*. 1970;72:7–21.

Chehade MJ, Pohl AP, Pearcy MJ, Nawana N. Clinical implications of stiffness and strength changes in fracture healing. *J Bone Joint Surg Br*. The British Editorial Society of Bone and Joint Surgery; 1997;79-B(1):9–12.

Chehade MJ, Solomon LB, Callary SA, Benveniste SH, Pohl AP, Howie DW. Differentially Loaded Radiostereometric Analysis to Monitor Fracture Stiffness: A Feasibility Study. *Clin Orthop Relat Res*. 2009;467(7):1839–47.

Choo AMT, Oxland TR. Improved RSA accuracy with DLT and balanced calibration marker distributions with an assessment of initial-calibration. *J Biomech*. Elsevier; 2003;36(2):259–64.

Christiansen J, Ejaz A, Nielsen PT, Laursen M. An Ultra-Short Femoral Neck-Preserving Hip Prosthesis: A 2-Year Follow-up Study with Radiostereometric Analysis and Dual X-Ray Absorptiometry in a Stepwise Introduction. *J Bone Joint Surg Am*. United States; 2020;102(2):128–36.

Claes LE, Cunningham JL. Monitoring the mechanical properties of healing bone. *Clin Orthop Relat Res*. Association of Bone and Joint Surgeons; 2009;467(8):1964–71.

Cnaan A, Laird NM, Slasor P. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat Med*. John Wiley & Sons, Ltd; 1997;16(20):2349–80.

Corrales LA, Morshed S, Bhandari M, Miclau T, III. Variability in the assessment of fracture-healing in orthopaedic trauma studies. *J Bone Joint Surg Am*. Journal of Bone and Joint Surgery, Inc.; 2008;90(9):1862–8.

Cyteval C. Imaging of knee implants and related complications. *Diagn Interv Imaging*. Elsevier Masson; 2016;97(7–8):809–21.

Denham RA. Hip mechanics. *J Bone Joint Surg Br*. J Bone Joint Surg Br; 1959;41 B:550–7.

Derbyshire B, Prescott RJ, Porter ML. Notes on the use and interpretation of radiostereometric analysis. *Acta Orthop*. 2009;80(1):124–30.

Dexter F, Shafer SL. Narrative Review of Statistical Reporting Checklists, Mandatory Statistical Editing, and Rectifying Common Problems in the Reporting of Scientific Articles. *Anesth Analg*. Lippincott Williams and Wilkins; 2017;124(3):943–7.

Digas G, Johansson PE, Kärrholm J. Inducible displacements of the cup and the femoral head during active range of motion: Dynamic RSA studies of cemented total hip replacements. *J Orthop Res*. J Orthop Res; 2013;31(11):1686–93.

Digas G, Thanner J, Anderberg C, Kärrholm J. Bioactive cement or ceramic/porous coating vs. conventional cement to obtain early stability of the acetabular cup. Randomised study of 96 hips followed with radiostereometry. J. Orthop. Res. United States: J Orthop Res; Sep, 2004 p. 1035–43.

Dmitrienko A, D'Agostino RB. Multiplicity Considerations in Clinical Trials. Longo DL, editor. *N Engl J Med*. 2018;378(22):2115–22.

Dodge Y. Cause and Effect in Epidemiology. *Concise Encycl Stat*. New York, NY: Springer New York; 2008. p. 62–4.

Downing MR, Ashcroft PB, Johnstone AJ, Bach O, Mackenzie S, Ashcroft GP. Assessment of inducible fracture micromotion in distal radial fractures using radiostereometry. *J Orthop Trauma*. 2008;22(8 Suppl):S96-105.

Dunbar MJ, Fong JW, Wilson DA, Hennigar AW, Francis PA, Glazebrook MA. Longitudinal migration and inducible displacement of the Mobility Total Ankle System. *Acta Orthop*. Acta Orthop; 2012;83(4):394–400.

Dutton A, Rubash HE. Hot topics and controversies in arthroplasty: cementless femoral fixation in elderly patients. *Instr Course Lect*. 2008;57:255–9.

Dyreborg K, Andersen M, Winther N, Solgaard S, Flivik G, Petersen MM. Migration of the uncemented Echo Bi-Metric and Bi-Metric THA stems: a randomized controlled RSA study involving 62 patients with 24-month follow-up. *Acta Orthop*. England; 2020;1–6.

Ebbinghaus S, Mjöberg B. Posterior angulation in trochanteric fractures detected with roentgen stereophotogrammetry. *Ups J Med Sci*. England; 1991;96(3):235–7.

Elmerson S, Andersson GBJ, Pope MH, Zetterberg C. Stability of fixation in femoral neck fractures: Comparison of four fixation devices in vivo and in cadavers. *Acta Orthop*. Informa Healthcare; 1987;58(2):109–12.

van Embden D, Stollenwerck GANL, Koster LA, Kaptein BL, Nelissen RGHH, Schipper IB. The stability of fixation of proximal femoral fractures: a radiostereometric analysis. *Bone Joint J*. England; 2015;97(3):391–7.

Engh CA, Massin P, Suthers KE. Roentgenographic assessment of the biologic fixation of porous-surfaced femoral components. *Clin Orthop Relat Res*. 1990;257(257):107–28.

Euler L. Formulae generales pro translatione quacunque corporum rigidorum. *Novi Comment Acad Sci Petropolitanae*. 1776;189–207.

FAR. Finnish Arthroplasty Register. https://www.thl.fi/far/.

Ferguson RJ, Broomfield JA, Malak TT, Palmer AJR, Whitwell D, Kendrick B, Taylor A, Glyn-Jones S. Primary stability of a short bone-conserving femoral stem: A two-year randomized controlled trial using radiostereometric analysis. *Bone Jt J*. British Editorial Society of Bone and Joint Surgery; 2018;100B(9):1148–56.

Fernández-Fernández R, Cruz-Pardos A, García-Rey E. Revision Total Hip Arthroplasty: Epidemiology and Causes. *Revis Total Jt Arthroplast*. Springer, Cham; 2020. p. 43–57.

Finch WH. Missing Data and Multiple Imputation in the Context of Multivariate Analysis of Variance. *J Exp Educ*. Routledge; 2016;84(2):356–72.

Finnilä S, Moritz N, Strandberg N, Alm JJ, Aro HT. Radiostereometric analysis of the initial stability of internally fixed femoral neck fractures under differential loading. *J Orthop Res*. 2019;37(1):239–47.

Finnilä S, Moritz N, Svedström E, Alm J, Aro H. Increased migration of uncemented acetabular cups in female total hip arthroplasty patients with low systemic bone mineral density. A 2-year RSA and 8-year radiographic follow-up study of 34 patients. *Acta Orthop*. 2016;87(1):48–54.

Floerkemeier T, Budde S, GV L, Windhagen H, HurSchler C, Schwarze M. Greater early migration of a short-stem total hip arthroplasty is not associated with an increased risk of osseointegration failure: 5th-year results from a prospective RSA study with 39 patients, a follow-up study. *Acta Orthop*. England; 2020;91(3):266–71.

Floerkemeier T, Schwarze M, Hurschler C, Gronewold J, Windhagen H, G von L, Budde S. The Influence of Tribological Pairings and Other Factors on Migration Patterns of Short Stems in Total Hip Arthroplasty. Biomed Res. Int. 2017. p. 8756432.

Flugsrud GB, Nordsletten L, Espehaug B, Havelin LI, Meyer HE. The effect of middle-age body weight and physical activity on the risk of early revision hip arthroplasty: A cohort study of 1,535 individuals. *Acta Orthop*. Taylor and Francis Ltd.; 2007;78(1):99–107.

Frazer AR, Tanzer M. New Implant Introduction in Total Hip Arthroplasty Using Radiostereometric Analysis: A Cautionary Note. *J Arthroplasty*. Churchill Livingstone Inc.; 2020;35(3):643–6.

Friedl G, Radl R, Stihsen C, Rehak P, Aigner R, Windhager R. The Effect of a Single Infusion of Zoledronic Acid on Early Implant Migration in Total Hip Arthroplasty. *J Bone Jt Surgery-American Vol*. Journal of Bone and Joint Surgery Inc.; 2009;91(2):274–81.

Galea VP, Botros MA, McTague MF, Weaver MJ, Vrahas MS, Malchau H, Heng M, Bragdon C. Radiostereometric analysis of stability and inducible micromotion after locked lateral plating of distal femur fractures. *J Orthop Trauma*. United States: Lippincott Williams and Wilkins; 2020;34(2):E60–6.

Garling EH, Kaptein BL, Geleijns K, Nelissen RG, Valstar ER. Marker Configuration Model-Based Roentgen Fluoroscopic Analysis. *J Biomech*. 2005;38(4):893–901.

Gaston MS, Simpson AHRW. Inhibition of fracture healing. J. Bone Jt. Surg. - Ser. B. The British Editorial Society of Bone and Joint Surgery; 2007. p. 1553–60.

Giori NJ, Ryd L, Carter DR. Mechanical influences on tissue differentiation at bone-cement interfaces. *J Arthroplasty*. 1995;10(4):514–22.

Glowacki J, Hurwitz S, Thornhill TS, Kelly M, LeBoff MS. Osteoporosis and vitamin-D deficiency among postmenopausal women with osteoarthritis undergoing total hip arthroplasty. *J Bone Joint Surg Am*. 2003;85-A(12):2371–7.

Glyn-Jones S, Palmer AJR, Agricola R, Price AJ, Vincent TL, Weinans H, Carr AJ. Osteoarthritis. *Lancet*. Lancet Publishing Group; 2015. p. 376–87.

Glyn-Jones S, Polgár K, Hicks J, Murray DW, Gill HS. RSA-measured inducible micromotion and interface modeling with finite element methods. *Clin Orthop Relat Res*. Lippincott Williams and Wilkins; 2006. p. 98–104.

Goodman SB. The effects of micromotion and particulate materials on tissue differentiation: Bone chamber studies in rabbits. *Acta Orthop Scand*. Taylor & Francis; 1994;65(sup258):1–43.

Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. Elsevier Limited; 2002. p. 248–52.

Gruen TA, McNeice GM, Amstutz HC. "Modes of failure" of cemented stem-type femoral components: a radiographic analysis of loosening. *Clin Orthop Relat Res*. United States; 1979;(141):17–27.

Habeck C, Stern Y. Multivariate Data Analysis for Neuroimaging Data: Overview and Application to Alzheimer's Disease. Cell Biochem. Biophys. Springer; 2010. p. 53–67.

Hansen L, de Raedt S, Jørgensen PB, Mygind-Klavsen B, Kaptein B, Stilling M. Dynamic radiostereometric analysis for evaluation of hip joint pathomechanics. *J Exp Orthop*. Springer; 2017;4(1):20.

Hansen L, De Raedt S, Jørgensen PB, Mygind-Klavsen B, Kaptein B, Stilling M. Marker free model-based radiostereometric analysis for evaluation of hip joint kinematics: A validation study. *Bone Joint Res*. British Editorial Society of Bone and Joint Surgery; 2018;7(6):379–87.

Hansson U, Toksvig-Larsen S, Jorn LP, Ryd L. Mobile vs. fixed meniscal bearing in total knee replacement: A randomised radiostereometric study. *Knee*. Knee; 2005;12(6):414–8.

Hjorth M, Søballe K, Jakobsen SS, Lorenzen N, Mechlenburg I, Stilling M. No association between serum metal ions and implant fixation in large-head metal-on-metal total hip arthroplasty. *Acta Orthop*. 2014;85(4):355–62.

Hoornenborg D, Sierevelt I, Spuijbroek J, Cheung J, van der Vis HM, Beimers L, Haverkamp D. Does hydroxyapatite coating enhance ingrowth and improve longevity of a Zweymuller type stem? A double-blinded randomised RSA trial. *Hip Int*. United States; 2018;28(2):115–21.

Howie DW, Holubowycz OT, Callary SA, Robertson TS, Solomon LB. Highly Porous Tantalum Acetabular Components Without Ancillary Screws Have Similar Migration to Porous Titanium Acetabular Components With Screws at 2 Years: A Randomized Controlled Trial. *J Arthroplasty*. United States: Churchill Livingstone Inc.; 2020;35(10):2931–7.

Hurschler C, Seehaus F, Emmerich J, Kaptein BL, Windhagen H. Comparison of the model-based and marker-based roentgen stereophotogrammetry methods in a typical clinical setting. *J Arthroplasty*. 2009;24(4):594–606.

Illgen R, Rubash HE. The optimal fixation of the cementless acetabular component in primary total hip arthroplasty. *J Am Acad Orthop Surg*. 2002;10(1):43–56.

Im G Il, Kim MK. The relationship between osteoarthritis and osteoporosis. J. Bone Miner. Metab. Springer-Verlag Tokyo; 2014. p. 101–9.

International Organization For Standardization. ISO 5725-1:1994 Accuracy (Trueness and Precision) of Measurement Methods and Results-Part 1: General Principles and Definitions. International Organization for Standardization Geneva, Switzerland; 1994.

International Organization For Standardization. ISO 16087:2013 Implants for surgery — Roentgen stereophotogrammetric analysis for the assessment of migration of orthopaedic implants. 2013.

Jacobsen A, Seehaus F, Hong Y, Cao H, Schuh A, Forst R, Sesselmann S. Model-based roentgen stereophotogrammetric analysis using elementary geometrical shape models: 10 years results of an uncemented acetabular cup component. *BMC Musculoskelet Disord*. BioMed Central Ltd.; 2018;19(1).

Jasty M, Bragdon C, Burke D, O'Connor D, Lowenstein J, Harris WH. In vivo skeletal responses to porous-surfaced implants subjected to small induced motions. *J Bone Joint Surg Am*. 1997;79(5):707–14.

Jiang J. Linear and Generalized Linear Mixed Models and Their Applications. Linear Gen. Linear Mix. Model. Their Appl. Springer New York; 2007.

Jones LC, Hungerford DS. Cement disease. *Clin Orthop Relat Res*. United States; 1987;(225):192–206.

Jørgensen PB, Daugaard H, Jakobsen SS, Lamm M, Søballe K, Stilling M, PB J, H D, SS J, M L, K S, M S. Higher early proximal migration of hemispherical cups with electrochemically applied hydroxyapatite (BoneMaster) on a porous surface compared with porous surface alone: a randomized RSA study with 53 patients. *Acta Orthop*. Taylor and Francis Ltd; 2020;91(1):26–32.

Kaptein BL, Valstar ER, Spoor CW, Stoel BC, Rozing PM. Model-based RSA of a Femoral Hip Stem Using Surface and Geometrical Shape Models. *Clin Orthop Relat Res*. 2006;(448):92–7.

Kaptein BL, Valstar ER, Stoel BC, Rozing PM, Reiber JHC. A new model-based RSA method validated using CAD models and models from reversed engineering. *J Biomech*. Elsevier; 2003;36(6):873–82.

Kärrholm J. Roentgen stereophotogrammetry: Review of orthopedic applications. *Acta Orthop Scand*. Taylor & Francis; 1989;60(4):491–503.

Kärrholm J, Gill RHS, Valstar ER. The history and future of radiostereometric analysis. *Clin Orthop Relat Res*. 2006;(448):10–21.

Kärrholm J, Herberts P, Hultmark P, Malchau H, Nivbrant B, Thanner J. Radiostereometry of hip prostheses. Review of methodology and clinical results. *Clin Orthop Relat Res*. 1997;(344):94–110.

Kärrholm J, Snorrason F. Migration of porous coated acetabular prostheses fixed with screws: Roentgen stereophotogrammetric analysis. *J Orthop Res*. 1992;10(6):826–35.

Konan S, Abdel MP, Haddad FS. Cemented versus uncemented hip implant fixation: Should there be age thresholds? *Bone Joint Res*. British Editorial Society of Bone and Joint Surgery; 2019;8(12):604.

Kooistra BW, Sprague S, Bhandari M, Schemitsch EH. Outcomes Assessment in Fracture Healing Trials: A Primer. *J Orthop Trauma*. 2010;24:S71–5.

Kopylov P, Aspenberg P, Yuan X, Ryd L. Radiostereometric analysis of distal radial fracture displacement during treatment: A randomized study comparing Norian SRS and external fixation in 23 patients. *Acta Orthop Scand*. England; 2001;72(1):57–61.

Krismer M, Bauer R, Tschupik J, Mayrhofer P. EBRA: A method to measure migration of acetabular components. *J Biomech*. Elsevier; 1995;28(10):1225–36.

Kundi M. Causality and the interpretation of epidemiologic evidence. Environ. Health Perspect. Public Health Services, US Dept of Health and Human Services; 2006. p. 969–74.

Ladurner A, Callary SA, Mitra A, Rickman M, Thewlis D, LB S. Radiostereometric Analysis Allows Assessment of the Stability and Inducible Displacement of Pelvic Ring Disruptions during Healing: A Case Series. *J Clin Med*. 2020;9(11).

Laende EK, Deluzio KJ, Hennigar AW, Dunbar MJ. Implementation and validation of an implant-based coordinate system for RSA migration calculation. *J Biomech*. Elsevier; 2009;42(14):2387–93.

Laende EK, Richardson CG, Dunbar MJ. A randomized controlled trial of tibial component migration with kinematic alignment using patient-specific instrumentation versus mechanical alignment using computer-assisted surgery in total knee arthroplasty. *Bone Jt J*. British Editorial Society of Bone and Joint Surgery; 2019a;101 B(8):929–40.

Laende EK, Richardson CG, Dunbar MJ. Predictive value of short-term migration in determining long-term stable fixation in cemented and cementless total knee arthroplasties. *Bone Joint J*. British Editorial Society of Bone and Joint Surgery; 2019b;101-B(7_Supple_C):55–60.

Laende EK, Richardson CG, Dunbar MJ. Migration and Wear of a Dual Mobility Acetabular Construct at 3 Years Measured by Radiostereometric Analysis. *J Arthroplasty*. United States; 2020;35(4):1109–16.

Lam Tin Cheung K, Lanting BA, McCalden RW, Yuan X, MacDonald SJ, Naudie DD, Teeter MG. Inducible displacement of cemented tibial components ten years after total knee arthroplasty. *Bone Jt J*. British Editorial Society of Bone and Joint Surgery; 2018;100B(2):170–5.

Lazarinis S, Milbrink J, Mattsson P, Mallmin H, Hailer NP. Bone loss around a stable, partly threaded hydroxyapatite-coated cup: A prospective cohort study using RSA and DXA. *HIP Int*. Wichtig Editore s.r.l.; 2014;24(2):155–66.

Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. Lancet. Elsevier; 2007. p. 1508–19.

Lindgren L, Jørgensen PB, Mørup RMS, Jensen M, Rømer L, Kaptein B, Stilling M. Similar patient positioning: A key factor in follow-up studies when using model-based radiostereometric analysis of the hip. *Radiography*. England: W.B. Saunders Ltd; 2020;26(2):e45–51.

Lingard EA, Mitchell SY, Francis RM, Rawlings D, Peaston R, Birrell FN, McCaskie AW. The prevalence of osteoporosis in patients with severe hip and knee osteoarthritis awaiting joint arthroplasty. *Age Ageing*. Age Ageing; 2009;39(2):234–9.

Little RB, Wevers HW, Siu D, Cooke TD V. A Three-Dimensional Finite Element Analysis of the Upper Tibia. *J Biomech Eng*. American Society of Mechanical Engineers; 1986;108(2):111.

Lord GA, Hardy JR, Kummer FJ. An uncemented total hip replacement: experimental study and review of 300 madreporique arthroplasties. *Clin Orthop Relat Res*. 1979;(141):2–16.

Madanat R, Mäkinen TJ, Aro HT, Bragdon C, Malchau H, Makinen TJ, Aro HT, Bragdon C, Malchau H. Adherence of hip and knee arthroplasty studies to RSA standardization guidelines. A systematic review. *Acta Orthop*. Taylor & Francis; 2014;85(5):447–55.

Madanat R, Moritz N, Aro HT. Three-dimensional computer simulation of radiostereometric analysis (RSA) in distal radius fractures. *J Biomech*. Elsevier; 2007;40(8):1855–61.

Madanat R, Moritz N, Larsson S, Aro HT. RSA Applications in Monitoring of Fracture Healing in Clinical Trials. *Scand J Surg*. SAGE PublicationsSage UK: London, England; 2006;95(2):119–27.

Madanat R, Strandberg N, Moritz N, Mattila K, Vahlberg T, Aro HT. Radiostereometric analysis in measurements of migration and inducible micromotion in intra-articular distal radius fractures treated with a volar plate. *J Orthop Trauma*. 2012;26(9):e153-60.

Mäkelä KT, Eskelinen A, Paavolainen P, Pulkkinen P, Remes V. Cementless total hip arthroplasty for primary osteoarthritis in patients aged 55 years and older: Results of the 8 most common cementless designs compared to cemented reference implants in the Finnish Arthroplasty Register. *Acta Orthop*. Taylor & Francis; 2010;81(1):42–52.

Mäkinen TJ, Alm JJ, Laine H, Svedström E, Aro HT. The incidence of osteopenia and osteoporosis in women with hip osteoarthritis scheduled for cementless total joint replacement. *Bone*. 2007;40(4):1041–7.

Mäkinen TJ, Koort JK, Mattila KT, Aro HT. Precision measurements of the RSA method using a phantom model of hip prosthesis. *J Biomech*. 2004;37(4).

Mäkinen TJ, Mattila KT, Määttänen H, Aro HT. Comparison of Digital and Conventional Radiostereometric Image Analysis in an Ankle Phantom Model. *Scand J Surg*. SAGE PublicationsSage UK: London, England; 2005;94(3):233–8.

Malak TT, Broomfield JAJ, Palmer AJR, Hopewell S, Carr A, Brown C, Prieto-Alhambra D, Glyn-Jones S. Surrogate markers of long-term outcome in primary total hip arthroplasty: A systematic review. *Bone Joint Res*. British Editorial Society of Bone and Joint Surgery; 2016;5(6):206–14.

Malchau H, Kärrholm J, Wang YX, Herberts P. Accuracy of migration analysis in hip arthroplasty - Digitized and conventional radiography, compared to radiostereometry in 51 patients. *Acta Orthop Scand*. Taylor & Francis; 1995;66(5):418–24.

Mann KA, Miller MA, Costa PA, Race A, Izant TH. Interface micromotion of uncemented femoral components from postmortem retrieved total hip replacements. *J Arthroplasty*. NIH Public Access; 2012;27(2):238-245.e1.

Marsell R, Einhorn TA. The biology of fracture healing. *Injury*. NIH Public Access; 2011;42(6):551–5.

Marsh DR, Li G. The biology of fracture healing: optimising outcome. *Br Med Bull*. Royal Society of Medicine Press Ltd; 1999;55(4):856–69.

Mattsson P, Larsson S. Stability of internally fixed femoral neck fractures augmented with resorbable cement: A prospective randomized study using radiostereometry. *Scand J Surg*.; 2003;92(3):215–9.

Mattsson P, Larsson S. Unstable Trochanteric Fractures Augmented with Calcium Phosphate Cement. *Scand J Surg*. SAGE PublicationsSage UK: London, England; 2004;93(3):223–8.

Maxwell SE, Kelley K. Ethics and sample size planning. *Handb ethics Quant Methodol*. New York, NY, US: Routledge/Taylor & Francis Group; 2011. p. 159–84.

McClelland D, Thomas PBM, Bancroft G, Moorcroft CI. Fracture Healing Assessment Comparing Stiffness Measurements using Radiographs. *Clin Orthop Relat Res*. 2007;(457):214–9.

McKee GK, Watson-Farrar J. Replacement of arthritic hips by the McKee-Farrar prosthesis. *J Bone Joint Surg Br*. England; 1966;48(2):245–59.

Mellon SJ, Liddle AD, Pandit H. Hip replacement: Landmark surgery in modern medical history. *Maturitas*. Elsevier; 2013;75(3):221–6.

Michelson JD, Riley LH. Considerations in the comparison of cemented and cementless total hip prostheses. *J Arthroplasty*. J Arthroplasty; 1989;4(4):327–34.

Minten MJM, Heesterbeek PJC, Spruit M, MJ M, PJ H, M S. No effect of additional screw fixation of a cementless, all-polyethylene press-fit socket on migration, wear, and clinical outcome. *Acta Orthop*. Taylor and Francis Ltd; 2016;87(4):363–7.

Miura M, Nakamura J, Matsuura Y, Wako Y, Suzuki T, Hagiwara S, Orita S, Inage K, Kawarai Y, Sugano M, Nawata K, Ohtori S. Prediction of fracture load and stiffness of the proximal femur by CT-based specimen specific finite element analysis: Cadaveric validation study. *BMC Musculoskelet Disord*. BioMed Central; 2017;18(1).

Mjöberg B. Is early migration enough to explain late clinical loosening of hip prostheses? *EFORT Open Rev*. British Editorial Society of Bone and Joint Surgery; 2020;5(2):113–7.

Mjöberg B, Hansson LI, Selvik G. Instability of total hip prostheses at rotational stress: A roentgen stereophotogrammetric study. *Acta Orthop Scand*. 1984;55(5):504–6.

Mogensen B, Ekelund L, Hansson LI, Lidgren L, Selvik G. Surface Replacement of the Hip in Chronic Arthritis: A Clinical, Radiographic and Roentgen Stereophotogrammetric Evaluation. *Acta Orthop Scand*. 1982;53(6):929–36.

Mohaddes M, Herberts P, Malchau H, PE J, Kärrholm J. High proximal migration in cemented acetabular revisions operated with bone impaction grafting; 47 revision cups followed with RSA for 17 years. *Hip Int*. United States; 2017;27(3):251–8.

Morshed S. Current Options for Determining Fracture Union. *Adv Med*. Hindawi Limited; 2014;2014:708574.

Morshed S, Corrales L, Genant H, Miclau T. Outcome Assessment in Clinical Trials of Fracture-Healing. *J Bone Jt Surgery-American Vol*. 2008;90(Suppl 1):62–7.

Moskal JT, Capps SG, Scanelli JA. Still no single gold standard for using cementless femoral stems routinely in total hip arthroplasty. *Arthroplast Today*. Elsevier; 2016;2(4):211–8.

Muharemovic O, Troelsen A, Thomsen MG, Kallemose T, Gosvig KK. A pilot study to determine the effect of radiographer training on radiostereometric analysis imaging technique. *Radiography*. W.B. Saunders; 2018;24(2):e37–43.

Munzinger U, Guggi T, Kaptein B, Persoon M, Valstar E, HC D. A titanium plasma-sprayed cup with and without hydroxyapatite-coating: a randomised radiostereometric study of stability and osseointegration. *Hip Int*. United States; 2013;23(1):33–9.

Murphy LB, Helmick CG, Schwartz TA, Renner JB, Tudor G, Koch GG, Dragomir AD, Kalsbeek WD, Luta G, Jordan JM. One in four people may develop symptomatic hip osteoarthritis in his or her lifetime. *Osteoarthr Cartil*. Osteoarthritis Cartilage; 2010;18(11):1372–9.

Naudie DDR, Somerville L, Korczak A, Yuan X, McCalden RW, Holdsworth D, Bourne RB. A randomized trial comparing acetabular component fixation of two porous ingrowth surfaces using RSA. *J Arthroplasty*. United States: Churchill Livingstone Inc.; 2013;28(8):48–52.

Nauth A, Creek AT, Zellar A, Lawendy AR, Dowrick A, Gupta A, Dadi A, van Kampen A, Yee A, de Vries AC, de Mol van Otterloo A, Garibaldi A, Liew A, McIntyre AW, Prasad AS, Romero AW, Rangan A, Oatt A, Sanghavi A, Foley AL, Karlsten A, Dolenc A, Bucknill A, Chia A, Evans A, Gong A, Schmidt AH, Marcantonio AJ, Jennings A, Ward A, Khanna A, Rai A, Smits AB, Horan AD, Brekke AC, Flynn A, Duraikannan A, Stødle A, van Vugt AB, Luther A, Zurcher AW, Jain A, Amundsen A, Moaveni A, Carr A, Sharma A, Hill AD, Trommer A, Rai BS, Hileman B, Schreurs B, Verhoeven B, Barden BB, Flatøy B, Cleffken BI, Bøe B, Perey B, Hanusch BC, Weening B, Fioole B, Rijbroek B, Crist BD, Halliday B, Peterson B, Mullis B, Richardson CG, Clark C, Sagebien CA, van der Pol CC, Bowler C, Humphrey CA, Coady C, Koppert CL, Coles C, Tannoury C, DePaolo CJ, Gayton C, Herriott C, Reeves C, Tieszer C, Dobb C, Anderson CG, Sage C, Cuento C, Jones CB, Bosman CHR, Linehan C, van der Hart CP, Henderson C, Lewis CG, Davis CA, Donohue C, Mauffrey C, Sundaresh DC, Farrell DJ, Whelan DB, Horwitz D, Stinner D, et al. Fracture fixation in the operative management of hip fractures (FAITH): an international, multicentre, randomised controlled trial. *Lancet*. Lancet Publishing Group; 2017;389(10078):1519–27.

Nazari-Farsani S, Vuopio M, Löyttyniemi E, Aro HT. Contributing factors to the initial femoral stem migration in cementless total hip arthroplasty of postmenopausal women. *J Biomech*. Elsevier Ltd; 2021;117:110262.

Nazari-Farsani S, Vuopio ME, Aro HT. Bone Mineral Density and Cortical-Bone Thickness of the Distal Radius Predict Femoral Stem Subsidence in Postmenopausal Women. *J Arthroplasty*. United States: Churchill Livingstone Inc.; 2020;35(7):1877-1884.e1.

Nelissen RGHH, Pijls BG, Kärrholm J, Malchau H, Nieuwenhuijse MJ, Valstar ER. RSA and Registries: The Quest for Phased Introduction of New Implants. *J Bone Jt Surgery-American Vol*. 2011;93(Suppl 3):62–5.

Niculescu M, Solomon BL, Viscopoleanu G, Antoniac IV. Evolution of Cementation Techniques and Bone Cements in Hip Arthroplasty. *Handb Bioceram Biocomposites*. Springer, Cham; 2016;859–99.

Nieuwenhuijse M, Vehmeijer S, Mathijsen N, Keizer S. Fixation of the short global tissue-sparing hip stem. *Bone Joint J*. England; 2020;102(6):699–708.

Nieuwenhuijse MJ, Valstar ER, Kaptein BL, Nelissen RGHH. The Exeter femoral stem continues to migrate during its first decade after implantation: 10-12 years of follow-up with radiostereometric analysis (RSA). *Acta Orthop*. Taylor & Francis; 2012. p. 129–34.

Nilsson KG, Theodoulou A, Mercer G, Quinn S, Krishnan J. Mid-term migration of a cementless, porous acetabular cup: A 5 year Radiostereometric analysis. *J Orthop*. 2017;14(4):454–60.

Nivbrant B, Kärrholm J. Migration and wear of hydroxyapatite-coated press-fit cups in revision hip arthroplasty: a radiostereometric study. J. Arthroplasty. United States; 1997. p. 904–12.

Nivbrant B, Kärrholm J, Önsten I, Carlsson A, Snorrason F. Migration of porous press-fit cups in hip revision arthroplasty. A radiostereometric 2-year follow-up study of 60 hips. J. Arthroplasty. United States; 1996. p. 390–6.

Olivecrona H, Maguire GQ, Noz ME, Zeleznik MP, Kesteris U, Weidenhielm L, Weidenhielm L. A CT method for following patients with both prosthetic replacement and implanted tantalum beads: preliminary analysis with a pelvic model and in seven patients. *J Orthop Surg Res*. BioMed Central; 2016;11:27.

Olivecrona L, Olivecrona H, Weidenhielm L, Noz ME, Maguire GQ, Zeleznik MP. Model studies on acetabular component migration in total hip arthroplasty using CT and a semiautomated program for volume merging. *Acta radiol*. 2003;44(4):419–29.

Önsten I, Åkesson K, Obrant KJ. Micromotion of the acetabular component and periacetabular bone morphology. *Clin Orthop Relat Res*. 1995;(310):103–10.

Önsten I, Berzins A, Shott S, Sumner DR. Accuracy and precision of radiostereometric analysis in the measurement of THR femoral component translations: human and canine in vitro models. *J Orthop Res*. John Wiley & Sons, Ltd; 2001;19(6):1162–7.

Önsten I, Carlsson A, Besjakov J. Wear in uncemented porous and cemented polyethylene sockets: a randomised, radiostereometric study. J. Bone Joint Surg. Br. England; 1998. p. 345–50.

Önsten I, Carlsson AS. Cemented versus uncemented socket in hip arthroplasty: A radiostereometric study of 60 randomized hips followed for 2 years. *Acta Orthop*. Informa Healthcare; 1994;65(5):517–21.

Önsten I, Carlsson AS, Ohlin A, Nilsson JA. Migration of acetabular components, inserted with and without cement, one- stage bilateral hip arthroplasty. A controlled, randomized study using roentgenstereophotogrammetric analysis. *J Bone Jt Surg - Ser A*. Journal of Bone and Joint Surgery Inc.; 1994;76(2):185–94.

Otten V, Maguire GQ, Noz ME, Zeleznik MP, Nilsson KG, Olivecrona H. Are CT Scans a Satisfactory Substitute for the Follow-Up of RSA Migration Studies of Uncemented Cups? A Comparison of RSA Double Examinations and CT Datasets of 46 Total Hip Arthroplasties. *Biomed Res Int*. Hindawi Limited; 2017;2017:3681458.

Otten VTC, Crnalic S, Röhrl SM, Nivbrant B, Nilsson KG. Stability of Uncemented Cups - Long-Term Effect of Screws, Pegs and HA Coating: A 14-Year RSA Follow-Up of Total Hip Arthroplasty. *J Arthroplasty*. United States: Churchill Livingstone Inc.; 2016;31(1):156–61.

Pakvis D, Luites J, G van H, Spruit M. A cementless, elastic press-fit socket with and without screws. *Acta Orthop*. 2012;83(5):481–7.

Perelgut ME, Polus JS, Lanting BA, Teeter MG. The effect of femoral stem collar on implant migration and clinical outcomes following direct anterior approach total hip arthroplasty. *Bone Joint J*. England; 2020;102(12):1654–61.

Peters SAE, Bots ML, Den Ruijter HM, Palmer MK, Grobbee DE, Crouse JR, O'Leary DH, Evans GW, Raichlen JS, Moons KGM, Koffijberg H. Multiple imputation of missing repeated outcome measurements did not add to linear mixed-effects models. *J Clin Epidemiol*. Pergamon; 2012;65(6):686–95.

Petersen MM, Nielsen PT, Lebech A, Toksvig-Larsen S, Lund B. Preoperative bone mineral density of the proximal tibia and migration of the tibial component after uncemented total knee arthroplasty. *J Arthroplasty*. Churchill Livingstone Inc.; 1999;14(1):77–81.

Pijls BG, Nelissen RGHH. The era of phased introduction of new implants. *Bone Joint Res*. British Editorial Society of Bone and Joint Surgery; 2016;5(6):215–7.

Pijls BG, Nieuwenhuijse MJ, Fiocco M, Plevier JW, Middeldorp S, Nelissen RG, Valstar ER. Early proximal migration of cups is associated with late revision in THA: a systematic review and meta-

analysis of 26 RSA studies and 49 survivalstudies. *Acta Orthop*. Taylor & Francis; 2012a;83(6):583–91.

Pijls BG, Valstar ER, Nouta K-A, Plevier JW, Fiocco M, Middeldorp S, Nelissen RG. Early migration of tibial components is associated with late revision. *Acta Orthop*. 2012b;83(6):614–24.

Pivec R, Johnson AJ, Mears SC, Mont MA. Hip arthroplasty. *Lancet*. Elsevier B.V.; 2012. p. 1768–77.

Praet F Van, Mulier M. To cement or not to cement acetabular cups in total hip arthroplasty: a systematic review and re-evaluation. *SICOT-J*. EDP Sciences; 2019;5.

Prieto-Alhambra D, Javaid MK, Judge A, Murray D, Carr A, Cooper C, Arden NK. Association between bisphosphonate use and implant survival after primary total arthroplasty of the knee or hip: population based retrospective cohort study. *BMJ*. British Medical Journal Publishing Group; 2011;343:d7222.

Prins AH, Kaptein BL, Stoel BC, Nelissen RGHH, Reiber JHC, Valstar ER. Handling modular hip implants in model-based RSA: Combined stem–head models. *J Biomech*. Elsevier; 2008;41(14):2912–7.

Probe R, Ward R. Internal fixation of femoral neck fractures. *J Am Acad Orthop Surg*. American Association of Orthopaedic Surgeons; 2006;14(9):565–71.

Probst KJ, Herp A, Leidel W. Stereo-radiologische Lokalisationsdiagnostik von Schalen- und Totalendoprothesen des Hüftgelenkes. *Arch Orthop Trauma Surg*. Springer-Verlag; 1978;92(1):75–81.

Prokopetz JJ, Losina E, Bliss RL, Wright J, Baron JA, Katz JN. Risk factors for revision of primary total hip arthroplasty: a systematic review. *BMC Musculoskelet Disord*. BioMed Central; 2012;13(1):251.

Ragnarsson JI, Boquist L, Ekelund L, Kärrholm J. Instability and femoral head vitality in fractures of the femoral neck. *Clin Orthop Relat Res*. United States; 1993;(287):30–40.

Ragnarsson JI, Eliasson P, Kärrholm J, Lundström B. The accuracy of measurements of femoral neck fractures. Conventional radiography versus roentgen stereophotogrammetric analysis. *Acta Orthop Scand*. England; 1992;63(2):152–6.

Ragnarsson JI, Hansson LI, Kärrholm J. Stability of femoral neck fractures. A postoperative roentgen stereophotogrammetric analysis. *Acta Orthop Scand*. England; 1989;60(3):283–7.

Ragnarsson JI, Kärrholm J. Stability of femoral neck fracture. Roentgen stereophotogrammetry of 29 hook-pinned fractures. *Acta Orthop Scand*. England; 1991;62(3):201–7.

Ragnarsson JI, Kärrholm J. Factors influencing postoperative movement in displaced femoral neck fractures: evaluation by conventional radiography and stereoradiography. *J Orthop Trauma*. United States; 1992;6(2):152–8.

Ranganathan P, Pramesh C, Buyse M. Common pitfalls in statistical analysis: Clinical versus statistical significance. *Perspect Clin Res*. Medknow; 2015;6(3):169.

Ranstam J, Ryd L, Önsten I. Accurate accuracy assessment: Review of basic principles. *Acta Orthop Scand*. 2000;71(1):106–8.

Regnér L, Carlsson L, Kärrholm J, Herberts P. Tibial component fixation in porous- and hydroxyapatite-coated total knee arthroplasty: A radiostereometric evaluation of migration and inducible displacement after 5 years. *J Arthroplasty*. Churchill Livingstone Inc.; 2000;15(6):681–9.

Reiner T, Sonntag R, JP K, Clarius M, Jakubowitz E, Weiss S, Ewerbeck V, Merle C, Moradi B, Kinkel S, Gotterbarm T, Hagmann S. The Migration Pattern of a Cementless Hydroxyapatite-Coated Titanium Stem under Immediate Full Weight-Bearing-A Randomized Controlled Trial Using Model-Based RSA. *J Clin Med*. 2020;9(7).

Richardson CG, Laende EK, Gross M, Dunbar MJ. Prospective clinical study using radiostereometric analysis (RSA) to evaluate fixation of a modular cemented polished femoral stem. *HIP Int*. United States: SAGE Publications Ltd; 2019;1120700019881429.

Richardson J, Cunningham J, Goodship A, O'Connor B, Kenwright J. Measuring stiffness can define healing of tibial fractures. *J Bone Joint Surg Br*. The British Editorial Society of Bone and Joint Surgery; 1994;76-B(3):389–94.

Ring PA. Complete replacement arthroplasty of the hip by the ring prosthesis. *J Bone Joint Surg Br*. England; 1968;50(4):720–31.

Robertson DM, St. Pierre L, Chahal R. Preliminary observations of bone ingrowth into porous materials. *J Biomed Mater Res*. 1976;10(3):335–44.

Röhrl SM, Nivbrant B, Ström H, Nilsson KG. Effect of augmented cup fixation on stability, wear, and osteolysis. *J Arthroplasty*. Elsevier BV; 2004;19(8):962–71.

Rosenberg EI, Bass PF, Davidson RA. Arriving at correct conclusions: The importance of association, Causality, and clinical significance. *South Med J*. 2012;105(3):161–6.

Rubin DB. Inference and Missing Data. *Biometrika*. JSTOR; 1976;63(3):581.

Ryd L. Micromotion in knee arthroplasty. *Acta Orthop Scand*. 1986;57(sup220):3–80.

Ryd L, Carlsson L, Herberts P. Micromotion of a noncemented tibial component with screw fixation. An in vivo roentgen stereophotogrammetric study of the Miller-Galante prosthesis. *Clin Orthop Relat Res*. Clin Orthop Relat Res; 1993;(295):218–25.

Ryd L, Lindstrand A, Stenström A, Selvik G. Porous Coated Anatomic Tricompartmental Tibial Components: The Relationship Between Prosthetic Position and Micromotion. *Clin Orthop Relat Res*. 1990;251.

Ryd L, Toksvig-Larsen S. Stability of the elevated fragment in tibial plateau fractures. A radiographic stereophotogrammetric study of postoperative healing. Int. Orthop. Germany; 1994. p. 131–4.

Ryd L, Yuan X, Löfgren H. Methods for determining the accuracy of radiostereometric analysis (RSA). *Acta Orthop Scand*. Taylor & Francis; 2000;71(4):403–8.

Saari T, Digas G, Kärrholm J. Risedronate does not enhance fixation or BMD in revision cups: randomised study with three years follow-up. *Hip Int*. United States; 2014;24(1):49–55.

Sadoghi P, Liebensteiner M, Agreiter M, Leithner A, Böhler N, Labek G. Revision surgery after total joint arthroplasty: A complication-based analysis using worldwide arthroplasty registers. *J Arthroplasty*. Churchill Livingstone; 2013;28(8):1329–32.

Sandberg O, Tholén S, Carlsson S, Wretenberg P. The anatomical SP-CL stem demonstrates a non-progressing migration pattern in the first year: a low dose CT-based migration study in 20 patients. *Acta Orthop*. Taylor and Francis Ltd.; 2020;1–6.

Sandgren B, Skorpil M, Nowik P, Olivecrona H, Crafoord J, Weidenhielm L, Persson A. Assessment of wear and periacetabular osteolysis using dual energy computed tomography on a pig cadaver to identify the lowest acceptable radiation dose. *Bone Joint Res*. The British Editorial Society of Bone and Joint SurgeryLondon; 2016;5(7):307–13.

SAS Inst. Inc. The MIXED Procedure. 2. print. Cary, NC: SAS Inst. Inc; 2009. p. 3886–4086.

Scheerlinck T, Polfliet M, Deklerck R, Van Gompel G, Buls N, Vandemeulebroucke J. Development and validation of an automated and marker-free CT-based spatial analysis method (CTSA) for assessment of femoral hip implant migration In vitro accuracy and precision comparable to that of radiostereometric analysis (RSA). *Acta Orthop*. Taylor & Francis; 2016;87(2):139–45.

Schielzeth H, Dingemanse NJ, Nakagawa S, Westneat DF, Allegue H, Teplitsky C, Réale D, Dochtermann NA, Garamszegi LZ, Araya-Ajoy YG. Robustness of linear mixed-effects models to violations of distributional assumptions. Sutherland C, editor. *Methods Ecol Evol*. British Ecological Society; 2020;11(9):1141–52.

Schütz U, Decking J, Decking R, Puhl W. Assessment of femoral component migration in total hip arthroplasty: digital measurements compared to RSA. *Acta Orthop Belg*. 2005;71(1):65–75.

Schwarze M, Budde S, von Lewinski G, Windhagen H, Keller MC, Seehaus F, Hurschler C, Floerkemeier T. No effect of conventional vs. minimally invasive surgical approach on clinical outcome and migration of a short stem total hip prosthesis at 2-year follow-up: A randomized controlled study. *Clin Biomech*. M. Schwarze, Laboratory for Biomechanics and Biomaterials, Anna-von-Borries-Str. 1-7, Hannover, Germany; 2018;51:105–12.

Seehaus F, Emmerich J, Kaptein BL, Windhagen H, Hurschler C. Experimental analysis of model-based roentgen stereophotogrammetric analysis (MBRSA) on four typical prosthesis components. *J Biomech Eng*. American Society of Mechanical Engineers; 2009;131(4):041004.

Seehaus F, Emmerich J, Kaptein BL, Windhagen H, Hurschler C. Dependence of model-based RSA accuracy on higher and lower implant surface model quality. *Biomed Eng Online*. BioMed Central; 2013;12(1):32.

Seehaus F, Olender GD, Kaptein BL, Ostermeier S, Hurschler C. Markerless Roentgen Stereophotogrammetric Analysis for in vivo implant migration measurement using three dimensional surface models to represent bone. *J Biomech*. 2012;45(8):1540–5.

Seehaus F, Schwarze M, Flörkemeier T, Lewinski G von, Kaptein BL, Jakubowitz E, Hurschler C. Use of single-representative reverse-engineered surface-models for RSA does not affect measurement accuracy and precision. *J Orthop Res*. John Wiley & Sons, Ltd; 2016;34(5):903–10.

Selvik G. Roentgen stereophotogrammetry: a method for the study of the kinematics of the skeletal system. *Acta Orthop Scand*. Taylor & Francis; 1989;60(sup232):1–51.

Selvik G. Roentgen Stereophotogrammetric Analysis. *Acta radiol*. 1990;31(2):113–26.

Selvik G, Alberius P, Aronson AS. A roentgen stereophotogrammetric system. Construction, calibration and technical accuracy. *Acta Radiol Diagn (Stockh)*. 1983;24(4):343–52.

Shah RP, MacLean L, Paprosky WG, Sporer S. Routine Use of Radiostereometric Analysis in Elective Hip and Knee Arthroplasty Patients: Surgical Impact, Safety, and Bead Stability. *J Am Acad Orthop Surg*. 2018;26(8):e173–80.

Shareghi B, PE J, Kärrholm J. Clinical evaluation of model-based radiostereometric analysis to measure femoral head penetration and cup migration in four different cup designs. *J Orthop Res*. United States; 2017;35(4):760–7.

Shen Y, Zhang YH, Shen L. Postmenopausal women with osteoporosis and osteoarthritis show different microstructural characteristics of trabecular bone in proximal tibia using high-resolution magnetic resonance imaging at 3 tesla. *BMC Musculoskelet Disord*. BioMed Central; 2013;14(1):136.

Shi J, Liang G, Huang R, Liao L, Qin D. Effects of bisphosphonates in preventing periprosthetic bone loss following total hip arthroplasty: a systematic review and meta-analysis. *J Orthop Surg Res*. BioMed Central; 2018;13(1):225.

Shi J, Zhu W, Liang S, Li H, Li S. Cross-Linked Versus Conventional Polyethylene for Long-Term Clinical Outcomes After Total Hip Arthroplasty: A Systematic Review and Meta-Analysis. *https://doi.org/101080/0894193920191606370*. Taylor & Francis; 2019;34(3):307–17.

Singer JM, Rocha FMM, Nobre JS. Graphical Tools for Detecting Departures from Linear Mixed Model Assumptions and Some Remedial Measures. *Int Stat Rev*. International Statistical Institute; 2017;85(2):290–324.

Sjöstedt Å, Zetterberg C, Hansson T, Hult E, Ekström L. Bone mineral content and fixation strength of femoral neck fractures: A cadaver study. *Acta Orthop*. Informa Healthcare; 1994;65(2):161–5.

Sköldenberg OG, Salemyr MO, Bodén HS, Lundberg A, Ahl TE, Adolphson PY. A new uncemented hydroxyapatite-coated femoral component for the treatment of femoral neck fractures. *J Bone Joint Surg Br*. 2011;93-B(5):665–77.

Sköldenberg OG, Sjöö H, Kelly-Pettersson P, Bodén H, Eisler T, Stark A, Muren O. Good stability but high periprosthetic bone mineral loss and late-occurring periprosthetic fractures with use of uncemented tapered femoral stems in patients with a femoral neck fracture. *Acta Orthop*. Taylor & Francis; 2014;85(4):396–402.

Snorrason F, Kärrholm J. Primary migration of fully-threaded acetabular prostheses. A roentgen stereophotogrammetric analysis. *J Bone Jt Surg Br Vol*. 1990;72(4):647–52.

Söderkvist I, Wedin P-Å. Determining the movements of the skeleton using well-configured markers. *J Biomech*. Elsevier; 1993;26(12):1473–7.

Solomon LB, Boopalan PRJVC, Chakrabarty A, Callary SA. Can tibial plateau fractures be reduced and stabilised through an angiosome-sparing antero-lateral approach? *Injury*. Netherlands; 2014;45(4):766–74.

Solomon LB, Callary SA, Stevenson AW, McGee MA, Chehade MJ, Howie DW. Weight-bearing-induced displacement and migration over time of fracture fragments following split depression fractures of the lateral tibial plateau. *J Bone Joint Surg Br*. England: The British Editorial Society of Bone and Joint Surgery; 2011;93-B(6):817–23.

Spector M. Historical review of porous-coated implants. *J Arthroplasty*. United States; 1987;2(2):163–77.

Spotorno L, Grappiolo G, Mumenthaler A. The CLS system. In: Morscher EW, editor. *Endoprosthetics*. Springer Science & Business Media; 2012. p. 271–96.

Sundfeldt M, V Carlsson L, B Johansson C, Thomsen P, Gretzer C. Aseptic loosening, not only a question of wear: A review of different theories. *Acta Orthop*. Taylor & Francis; 2006;77(2):177–97.

Tabori-Jensen S, Mosegaard SB, Hansen TB, Stilling M. Inferior stabilization of cementless compared with cemented dual-mobility cups in elderly osteoarthrosis patients: a randomized controlled radiostereometry study on 60 patients with 2 years' follow-up. *Acta Orthop*. Taylor and Francis Ltd; 2020;91(3):246–53.

Tagare P. Signal averaging. *Biomed Digit signal Process*. 1993. p. 184–92.

Thanner J, Kärrholm J, Malchau H, Wallinder L, Herberts P. Migration of press-fit cups fixed with poly-L-lactic acid or titanium screws: a randomized study using radiostereometry. J. Orthop. Res. United States; 1996. p. 895–900.

Thewlis D, Callary SA, Fraysse F, Solomon LB. Peak loading during walking is not associated with fracture migration following tibial plateau fracture: A preliminary case series. *J Orthop Res*. United States: John Wiley & Sons, Ltd; 2015;33(9):1398–406.

Thien T, Ahnfelt L, Eriksson M, Strömberg C, Kärrholm J. Immediate weight bearing after uncemented total hip arthroplasty with an anteverted stem: a prospective randomized comparison using radiostereometry. *Acta Orthop*. England; 2007;78(6):730–8.

Troelsen A, Malchau E, Sillesen N, Malchau H. A Review of Current Fixation Use and Registry Outcomes in Total Hip Arthroplasty: The Uncemented Paradox. *Clin Orthop Relat Res*. 2013;471(7):2052–9.

Trozzi C, Kaptein BL, Garling EH, Shelyakova T, Russo A, Bragonzoni L, Martelli S. Precision assessment of model-based RSA for a total knee prosthesis in a biplanar set-up. *Knee*. Elsevier; 2008;15(5):396–402.

Ulrich SD, Seyler TM, Bennett D, Delanois RE, Saleh KJ, Thongtrangan I, Kuskowski M, Cheng EY, Sharkey PF, Parvizi J, Stiehl JB, Mont MA. Total hip arthroplasties: What are the reasons for revision? *Int Orthop*. Springer; 2008;32(5):597–604.

Valstar ER. Digital Roentgen Stereophotogrammetry. Development, validation, and clinical application. University of Leiden; 2001.

Valstar ER, Gill R, Ryd L, Flivik G, Borlin N, Kärrholm J, Börlin N, Kärrholm J. Guidelines for standardization of radiostereometry (RSA) of implants. *Acta Orthop*. 2005;76(4):563–72.

Valstar ER, de Jong FW, Vrooman HA, Rozing PM, Reiber JHC. Model-based Roentgen stereophotogrammetry of orthopaedic implants. *J Biomech*. Elsevier; 2001;34(6):715–22.

Valstar ER, Spoor CW, Nelissen RGHH, Rozing PM. Roentgen stereophotogrammetric analysis of metal-backed hemispherical cups without attached markers. *J Orthop Res*. John Wiley & Sons, Ltd; 1997;15(6):869–73.

Vanrusselt J, Vansevenant M, Vanderschueren G, Vanhoenacker F. Postoperative radiograph of the hip arthroplasty: what the radiologist should know. *Insights Imaging*. Springer; 2015;6(6):591–600.

Verbeke G, Fieuws S, Molenberghs G, Davidian M. The analysis of multivariate longitudinal data: A review. *Stat Methods Med Res*. NIH Public Access; 2014;23(1):42–9.

Verbeke G, Molenberghs G. Linear mixed models in practice: a SAS-oriented approach. Springer Science & Business Media; 2012.

Vollmer SH, Howard G. Statistical Power, the Belmont Report, and the Ethics of Clinical Trials. *Sci Eng Ethics*. Springer; 2010;16(4):675–91.

van der Voort P, Pijls BG, Nieuwenhuijse MJ, Jasper J, Fiocco M, Plevier JWM, Middeldorp S, Valstar ER, Nelissen RGHH. Early subsidence of shape-closed hip arthroplasty stems is associated with late revision. *Acta Orthop*. 2015;86(5):575–85.

de Vries LM, van der Weegen W, Pilot P, Stolarczyk PA, Sijbesma T, Hoffman EL. The predictive value of radiostereometric analysis for stem survival in total hip arthroplasty. A systematic review. *Hip Int*. 2014;24(3):215–22.

Weber M, Renkawitz T, Voellner F, Craiovan B, Greimel F, Worlicek M, Grifka J, Benditz A. Revision Surgery in Total Joint Replacement Is Cost-Intensive. *Biomed Res Int*. Hindawi Limited; 2018;2018:8987104.

West BT, Welch KB, Galecki AT. A Brief History of LMMs. *Linear Mix Model A Pract Guid Using Stat Softw*. 2nd ed. CRC Press; 2006. p. 5–8.

Wilkinson JM, Stockley I, Peel NFA, Hamer AJ, Elson RA, Barrington NA, Eastell R. Effect of pamidronate in preventing local bone loss after total hip arthroplasty: A randomized, double-blind, controlled trial. *J Bone Miner Res*. J Bone Miner Res; 2001;16(3):556–64.

Willert HG, Ludwig J, Semlitsch M. Reaction of bone to methacrylate after hip arthroplasty: a long-term gross, light microscopic, and scanning electron microscopic study. *J Bone Joint Surg Am*. United States; 1974;56(7):1368–82.

Wilson DAJ, Astephen JL, Hennigar AW, Dunbar MJ. Inducible Displacement of a Trabecular Metal Tibial Monoblock Component. *J Arthroplasty*. Churchill Livingstone; 2010;25(6):893–900.

Wolf O, Mattsson P, Milbrink J, Larsson S, Mallmin H. The effects of different weight-bearing regimes on press-fit cup stability: a randomised study with five years of follow-up using radiostereometry. *Int Orthop*. 2012;36(4):735–40.

Wolf O, Milbrink J, Larsson S, Mattsson P, Mallmin H. The optimal timing of baseline radiostereometric analysis of uncemented press fit cups. *Scand J Surg*. England: Finnish Surgical Society; 2010;99(4):244–9.

Wyatt MC, Roberton A, Foxall-Smi M, Beswick AD, Kunutsor SK, Whitehouse MR. Does vitamin E highly-crosslinked polyethylene convey an advantage in primary total hip replacement? A systematic review and meta-analysis: *https://doi.org/101177/1120700019858335*. SAGE PublicationsSage UK: London, England; 2019;30(5):598–608.

Yamada H, Yoshihara Y, Henmi O, Morita M, Shiromoto Y, Kawano T, Kanaji A, Ando K, Nakagawa M, Kosaki N, Fukaya E. Cementless total hip replacement: past, present, and future. *J Orthop Sci*. Springer; 2009;14(2):228.

Yu CH. A Model Must Be Wrong to be Useful: The Role of Linear Modeling and False Assumptions in Theoretical Explanation. *Open Stat Probab J*. 2010;2(1).

Zewotir T, Galpin JS. Influence diagnostics for linear mixed models. *J data Sci*. 2005;3(2):153–77.

Zhou ZK, Li MG, Börlin N, Wood DJ, Nivbrant B. No increased migration in cups with ceramic-on-ceramic bearing: An RSA study. *Clin Orthop Relat Res*. Lippincott Williams and Wilkins; 2006. p. 39–45.

Zielinski SM, Keijsers NL, Praet SFE, Heetveld MJ, Bhandari M, Wilssens JP, Patka P, Van Lieshout EMM. Femoral neck shortening after internal fixation of a femoral neck fracture. *Orthopedics*. 2013;36(7):e849–58.

Zimmerman DW. A note on the influence of outliers on parametric and nonparametric tests. *J Gen Psychol*. Taylor & Francis Group; 1994;121(4):391–401.

Zlowodzki M, Ayieni O, Petrisor BA, Bhandari M. Femoral neck shortening after fracture fixation with multiple cancellous screws: Incidence and effect on function. *J Trauma - Inj Infect Crit Care*. J Trauma; 2008;64(1):163–9.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2021

ISBN 978-951-29-8619-4 (PRINT)
ISBN 978-951-29-8620-0 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)