Nienke de Laat
OP3 SUMMA-tech
Department of Orthopedics
Leiden University Medical Centre

Accuracy and Precision of Volumetric Matching Micromotion Analysis (V3MA) for tibial component migration measurement in total knee arthroplasty

An experimental study and clinical proof of concept in six patients

N.N. de Laat

Leids Universitair Medisch Centrum

February 17th 2023

TABLE OF CONTENTS

ABSTRACT

Introduction: Early migration and continuous migration in the first two years is considered a predictive factor of aseptic loosening in total knee arthroplasty (TKA). The current gold standard to measure implant migration is roentgen stereophotogrammetric analysis (RSA). However, RSA is complex to implement in common practice due to the use of a calibration cage, trained radiology personnel, and insertion of tantalum markers in the bone. Therefore, computed tomography based migration analysis (CTBMA) has been proposed as an alternative method. CTBMA measures the displacement of implants relative to the host bone over time in CT images. We developed a volume-based and marker-free CTBMA called VoluMetric Matching Micromotion Analysis (V3MA). The objective of this study was to validate V3MA in tibial components by evaluating the accuracy and precision *in vitro* and to provide clinical proof of concept.

Methods: Three experiments (A, B, C) were conducted. In all experiments, tibial component migration was measured with both V3MA and RSA. (A) Accuracy: 21 different positions of a tibial component attached to a micromanipulator were assessed with respect to the cadaveric human tibial bone. (B) Precision: Repeated zero-migration measurements of 4 sets consisting of a cadaveric human tibial bone with cemented tibial component. (C) Clinical proof of concept: Clinical data of 6 patients. Data analysis included descriptive statistics (A, B, C), Bland-Altman analysis (A, C), and interclass correlation coefficient (ICC) (C).
The **primary outcome** was the accuracy (A) and precision (B) of both V3MA and RSA of tibial components *in vitro*. **Secondary outcomes** included agreement of clinical tibial migration measured with V3MA and RSA (C), as well as an explorative interobserver variability (C).

Results: (A) V3MA accuracy ($t * RMS$) ranged between 0.02 mm and 0.15 mm for translations and was 0.03° in Y-rotation. For RSA, the accuracy ranged between 0.09 mm and 0.33 mm for translations and was 0.25° in the Y-rotation. V3MA accuracy was significantly better in anteroposterior translation (p = 0.01). (B) V3MA precision (SD) ranged from 0.01 to 0.06 mm for translations and 0.02 to 0.07° in rotations. RSA precision ranged from 0.00 to 0.06 mm for translations and 0.04 to 0.25° in rotations. V3MA precision was significantly better in all directions, except superoinferior translation. (C) The mean difference (limits of agreement) between V3MA and RSA was -0.05 mm (-0.49 to 0.39) and -0.14° (-0.79 to 0.51) for total translation and total rotation, respectively. There was a good correlation between two observers using V3MA (ICC: 0.76, 95%CI: 0.58 to 0.87).

Discussion and conclusion: The accuracy and precision of V3MA were comparable to RSA. However, results may differ for other CT resolutions and in other joints. Furthermore, V3MA was feasible and comparable to RSA in clinical practice, with good reliability between observers. Therefore, V3MA seems a promising alternative to RSA in migration measurement of tibial components in TKA.

(464 words)

1. INTRODUCTION

Globally more than 1 million primary and revision total knee arthroplasties (TKA) are performed each year [1]. Approximately 82% of primary TKAs last 25 years [2]. Although the overall survival rate of knee implants is high, some implants do not perform as well as others. For example, 29% of the Accord knee implant (DePuy International Ltd., Leeds, United Kingdom) required revision within 12 years [3], and 13% of the St. Leger knee implant (Covision, Carlton in Lindrick, United Kingdom) was revised within 10 years [4]. The most reported reason for revision surgery was aseptic loosening of either the tibial or femoral component: 26 of 27 revised knees (Accord) and 12 of 19 revised knees (St. Leger). In large national registries, aseptic loosening is also one of the most common reasons for revision in TKA [5], [6]. Revision surgery is complex, expensive and often with less favorable outcomes. To improve primary TKA survival, new implants, fixation methods, and surgical techniques should be evaluated in a phased evidenced-based introduction, as proposed by several authors [7]–[11].

Early migration and continuous migration in the first two years are associated with late revision for aseptic loosening [12]–[14]. Consequently, implant migration measurement can be used as an objective measurement in the phased introduction of new implants. Currently, the gold standard for migration measurement is radiostereometric analysis (RSA), which was introduced by Selvik in 1974 [15], [16]. Nowadays, an enhanced version of RSA is used (model-based RSA) [17], [18]. RSA measures the migration, a shift in position over time, of the implant relative to the bone.

RSA has some disadvantages though. RSA requires a calibration cage, a special set-up of two roentgen tubes, trained radiology personnel, and the insertion of markers into the bone during surgery. In addition, these markers are sensitive to being superimposed by the implant (marker occlusion). To overcome these disadvantages, a method using regular computer tomography (CT) images has been proposed as an alternative to measure implant migration [19]–[21]. This CT-based migration analysis (CTBMA) would allow more hospitals to evaluate orthopedic implants because CT scanners are widely available compared to RSA research set-ups. Nevertheless, its accuracy[1] and precision[2] must be comparable to that of RSA.

In 2002 Olivecrona et al. introduced the first CTBMA to measure acetabular cup migration [22]. Thereafter, CTBMA was tested in acetabular cups [19], [20], [23], [24], shoulder implants [25] and in femoral hip implants [21], [26]. However, none of these studies have tested CTBMA in TKA.

Three different techniques were used in CTBMA: marker-based registration [20], marker-free surface registration [24], [25], and marker-free volume registration [21]. In volume registration the underlying voxel intensities of the selected volumes in the baseline image are matched to the follow-up image using digital image registration [27]. We developed a marker-free, volume-based CTBMA software called VoluMetric Matching Micromotion Analysis (V3MA).

This study aimed to validate V3MA for tibial components in TKA using *in vitro* experiments and provide clinical proof of concept in a clinical dataset. Therefore, three research questions were formulated: (1) What is the *in vitro* accuracy and precision of V3MA and RSA for tibial components? (2) What are the limits of agreement between V3MA and RSA, and between two observers using V3MA in 6 patients? (3) What is the interobserver variability between two observers operating with V3MA in 6 patients?

[1] *Accuracy* is defined as "the closeness of agreement between a measured quantity value and a true quantity value of a measurand" (ISO 16087:2013).

[2] *Precision* is defined as "the degree to which repeated measurements under unchanged condition show the same results" (ISO 16087:2013).

2. MATERIALS AND METHOD

The validation of V3MA in tibial components consisted of 3 experiments:

(A) *in vitro* accuracy of V3MA and RSA
(B) *in vitro* precision of V3MA and RSA
(C) clinical proof of concept using V3MA and RSA in 6 TKA patients; including interobserver variability

Data from each experiment (A, B, and C) was tested separately for normality using density plots and the Shapiro-Wilk test. Additional statistical tests were described for each experiment. All tests described in this document with a p-value < 0.05 were considered significant. All data was analyzed using R (version 3.6.1) within RStudio (version 1.1.456).

2.1. Output parameters

Each measurement in experiments A, B, and C included 3 translation parameters in millimeters (mm) and 3 rotation parameters in degrees (°). Thus, migration consisted of the following 6 parameters: $\boldsymbol{M} = [T_x, T_y, T_z, R_x, R_y, R_z]$. These migration parameters were used to calculate total translation (TT) and total rotation (TR) using the 3D Pythagorean theorem (Equation (1)).

$$TT = \sqrt{{T_x}^2 + {T_y}^2 + {T_z}^2}$$
$$TR = \sqrt{{R_x}^2 + {R_y}^2 + {R_z}^2} \qquad (\ 1\ )$$

Migration was measured with V3MA (section 2.5) and RSA (section 2.6). The coordinate system of V3MA equaled the coordinate system of the RSA set-up, defined by the calibration cage used. In the current study, right and left knees had the same coordinate system (Figure 1).

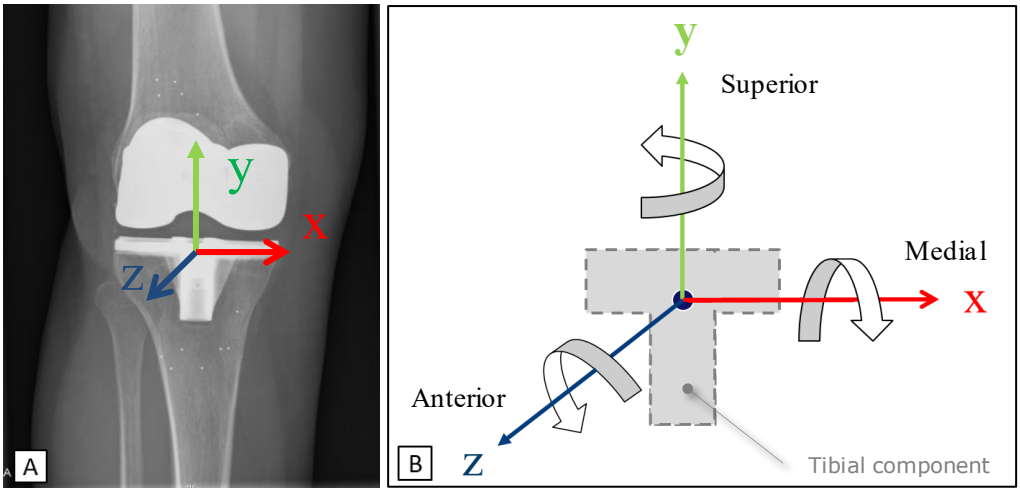


Figure 1: A) Schematic visualization of the coordinate system of RSA and V3MA with the center of the tibial component as origin. The z-axis points out of page/screen. B) Translation along the x-axis (Tx) is mediolateral, along the y-axis (Ty) is superoinferior, and along the z-axis (Tz) is anteroposterior. Rotations are anteroposterior around the x-axis (Rx), endo-exo around the y-axis (Ry), and mediolateral around the z-axis (Rz). Axes are defined for a right knee.

2.2. Experiment A (*in vitro* accuracy)

2.2.1. Set-up

The set-up consisted of a cadaveric human tibial bone and an uncemented tibial component (NexGen, size 5) attached to a micromanipulator (accuracy of micromanipulator: 0.02 mm / 0.01 degrees), which was assumed as the true value (Figure 2). The imposed migration by the micromanipulator

was measured with magnetic linear encoders (MRS130, IES Elektronics LTD, Esenler-Istanbul-Turkey) and was displayed digitally (DC-503, Ditron, Chengdu, Sichuan, China). A no-touching area between the implant and bone allowed free movement of the tibial component. Bone markers were attached to the tibial bone for RSA after acquisition of CT images. A total of 21 different positions were imposed by the micromanipulator from starting position, consisting of 5 translations along each axis (T_x, T_y, T_z with translations of 0.1; 0.2; 0.5; 1.0 and 1.5 mm) and 6 rotations around the y-axis (R_y with rotations of 0.5; 1 and 1.5 degrees in positive and negative direction). The micromanipulator could only rotate around the y-axis and not around the x- and z-axis. Before each set of translations or rotations, the micromanipulator was set at the zero position and a baseline image was taken according to the scan protocol (Appendix A). This protocol resulted in 26 CT scans and 26 paired RSA images: 5 baseline images + 21 images of the imposed positions. In addition to the different positions applied to the micromanipulator, the position of the entire set-up was changed between each CT scan and similarly between each RSA radiograph to mimic clinical positional variation.

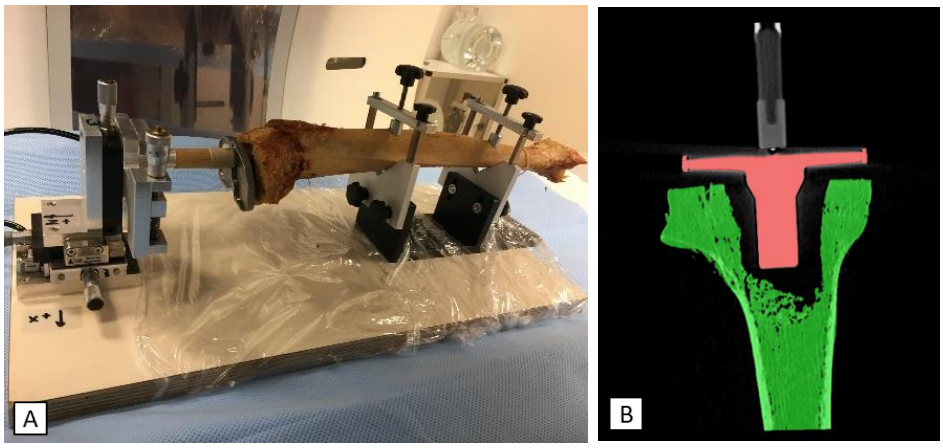


Figure 2: A) Set-up of experiment A to determine the accuracy. Including an uncemented tibial component attached to the micromanipulator, allowing free movement of the implant in controlled steps. The cadaveric human tibial bone was fixed to a wooden board. B) V3MA visualization of a coronal slice with tibial component mask (red) and bone mask (green).

2.2.2. Statistical analysis accuracy

For both V3MA and RSA the difference between migration result ($\boldsymbol{M}_r$) and displacement imposed by the micromanipulator ($\boldsymbol{M}_i$) was calculated by $\boldsymbol{\Delta} = \boldsymbol{M}_r - \boldsymbol{M}_i$ for each measurement (4 output parameters: T_x, T_y, T_z, R_y). The accuracy for each parameter was described by descriptive statistics (mean, SD, minimum, maximal) and $t * RMS$, where the RMS represented the root mean square error [24], [28] (Equation (*2*)).

$$RMS = \sqrt{\frac{{\Delta_1}^2 + {\Delta_2}^2 \dots + {\Delta_n}^2}{n}} \qquad (2)$$

Where n was the number of measurements and t represented the value obtained in a t-distribution table for a two-tailed t-distribution with a 95% confidence level with $(n-1)$ degrees of freedom. This resulted in $t = 2.776$ for translations (4 degrees of freedom) and $t = 2.571$ for rotations (5 degrees of freedom). A two-sided paired t-test was used to compare the means of V3MA and RSA. Levene's test was used to test for equality of variances in V3MA and RSA. Bland-Altman plots were constructed to visually show the agreement between imposed and measured values, as well as systematic bias and proportional bias [29]. Also, the mean difference and limits of agreement (mean ± $1.96 * SD$) between method (V3MA or RSA) and micromanipulator were determined.

2.3. Experiment B (*in vitro* precision)

2.3.1. Set-up

The set-up consisted of a cadaveric human tibial and femoral bone with cemented total knee implant (Figure 3). Only tibial migration was analyzed, however, we did include the femoral component in the CT scans to account for realistic metal artifacts. We used 4 sets of different cadaveric bones with cemented tibial components (set 1 and set 2: LPS FL ST mobile bearing size 5, set 3: LPS FL ST mobile bearing size 7 and set 4: unknown). All tibial components were firmly attached to the bone. Therefore, zero migration was expected. The position of the set-up was changed between each CT scan to mimic patient motion during consecutive outpatient clinic visits.

Each set was scanned 5 times in the CT scanner, of which the first was used as baseline and the other 4 as follow-up. This protocol resulted in a total of 20 CT volumes and 16 migration measurements. This experiment was repeated for RSA after attaching the bone markers to the tibial bone. Only the sets with known tibial components were usable for RSA. Therefore, migration could only be measured in 3 of the 4 sets resulting in 12 migration measurements with RSA.

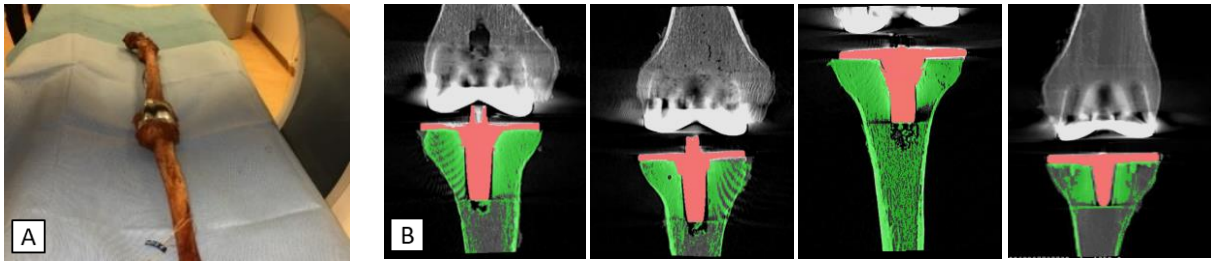


Figure 3: A) Set-up of experiment B to determine the precision. The tibial and femoral components were cemented in the cadaveric human bones, therefore zero migration was expected.
B) V3MA visualization of set 1 to set 4 (from left to right) with masks of the tibial component (red) and the tibia bone (green).

2.3.2. Statistical analysis precision

Descriptive statistics (mean, SD, minimum, maximal) and 95%-CI were calculated (8 output parameters: $T_x, T_y, T_z, R_x, R_y, R_z, TT, TR$). The 95%-CI was calculated by mean ± $t \cdot SD$ where t represented the value obtained in a t-distribution table used in small sample sizes [24], [28]. This resulted in $t = 2.131$ for V3MA (15 degrees of freedom) and $t = 2.201$ for RSA (11 degrees of freedom) for a double-sided 95%-CI. A two-sided paired t-test was used to compare the means between V3MA and RSA. Levene's test was used to test for equality of variances in V3MA and RSA.

2.4. Experiment C (clinical proof of concept and interobserver variability)

2.4.1. Clinical data

Six patients of a prospective randomized clinical trial (RCT) comparing 2 cemented tibial components (Persona PS and NexGen LPS) [30] were selected to provide clinical proof of concept and interobserver variability. The clinical data included CT scans and RSA images at 1 (baseline) and 5 (follow-up) years after TKA for each patient. Patients were selected based on the resolution of the CT images. The selection criteria was a pixel size smaller than 0.40 mm, which was similar to the resolution of experiment A and B. Only 6 cases met this criteria for both baseline and follow-up CT image. Tibial implant migration was measured with V3MA and RSA. RSA was performed by an expert in RSA (LK), and V3MA was performed by two observers (NL and BK) to determine interobserver variability. Both technicians semi-automatically segmented the tibial component and tibial bone, which were used for V3MA.

2.4.2. Statistical analysis

Bland-Altman plots were constructed to determine the mean differences and limits of agreement between V3MA (NL) and RSA, and between two observers operating with V3MA (NL and BK) (8 output parameters: $T_x, T_y, T_z, R_x, R_y, R_z, TT, TR$) [29]. Also, the interclass correlation coefficient (ICC) was calculated.

2.5. V3MA

The CT images were analyzed with a novel in-house developed software (V3MA, Python v3.9.7, version/date: 02-04-2022). V3MA measured the displacement of orthopedic implants relative to the host bone over time. In general, V3MA consists of 4 steps; (1) data acquisition, (2) defining masks, (3) registration, and (4) calculation of migration. Technical details were provided (Appendix B) [31].

1. CT images were acquired on a Toshiba Aquilion ONE CT scanner at Leiden University Medical Hospital (LUMC; The Netherlands). All images were reconstructed with a bone filter (convolution kernel FC30) without the use of metal artifact reduction algorithms. The slice thickness was 0.5 mm, and pixel sizes ranged between 0.30 mm and 0.38 mm (Table 1).
2. Defining masks of the baseline image was performed by semi-automatic segmentation with Mimics (Mimics, Materialise, Belgium). The segmentation included the following steps:
    a. A threshold derived from an intensity profile line.
    b. Region grow to include connecting voxels.
    c. Morphological operations (open and close) of 1 pixel to break small connections and to fill cavities within the mask (only in the mask of tibial component).
    d. A 3D model (.STL file) was constructed.

    Three masks were obtained by segmentation: an align mask (including bone, implant, and soft tissue if available), a tibial bone mask, and a tibial component mask. These three masks were used for the image registration.
3. Registration is the spatial alignment of two or more images and was performed in V3MA using Elastix [27]. Three registrations were performed using the voxel intensities within the three defined masks:
    a. Align registration: rough alignment of baseline and follow-up image using fewer iterations in each resolution level of the multiresolution approach.
    b. Implant-implant registration: align registration was used as starting point.
    c. Bone-bone registration: implant registration was used as starting point.

    All registrations were visually verified in V3MA by displaying the two images on top of each other with a checkerboard and red/green view. Bone and implant were considered rigid objects; therefore v3MA found rigid transformations between baseline and follow-up CT images.
4. Migration $\boldsymbol{M}$ was automatically calculated in V3MA by determining the relative movement of the two volumes (i.e., tibial component and tibia bone) with respect to a fixed migration coordinate system with its origin in the geometric center of the implant (Figure 1). Implant migration was calculated in terms of translations and rotations (section 2.1).

Table 1: CT parameters. The matrix size in all CT images is 512x512 (rows x columns). Abbreviations: CT = Computed Tomography, kVp = kilo volt, mA = milli ampere, mm = millimeter, mGy = milli gray.

CT parameter	Experiment A	Experiment B	Experiment C
CT scanner	Aquilion ONE	Aquilion ONE	Aquilion ONE
Helical / Volume	Volume	Volume	Helical and Volume
CT tube voltage (kVp)	120	120	120 to 135
Data collection diameter	500	500	320 to 500
CT tube current (mA)	30	30	66 to 110
Reconstruction diameter (mm)	154.296	160.156^	152.50 to 195.312
Convolution kernel	FC30†	FC30†	FC30†
Focal spot sizes (mm)	0.9\0.8	0.9\0.8	0.9\0.8
Slice thickness (mm)	0.50	0.50	0.50
Pixel spacing (mm)*	0.30	0.31^	0.30 to 0.38
CTDIvol (mGy)	2.3	2.3	3.0 to 7.6

**Pixel spacing is equal in both direction (square) and equals pixel size assuming the space between pixels is zero*
^for 5/20 CT volumes the reconstruction diameter was 500 resulting in a pixel size of 0.98 mm (set 4)
†FC30 refers to a Toshiba convolution kernel used in image reconstruction

2.6. RSA

RSA radiographs were acquired using two X-ray sources (Oldelft-Benelux, The Netherlands) angled at 40° towards each other in combination with digital radiography detectors (Canon CDXI, 434*35 cm; Oldelft-Benelux) and a uniplanar calibration cage (Medis CarbonBox 008; Medis Specials, The Netherlands) (Figure 4)[15]. RSA was performed by an RSA expert (LK) using Model-based RSA software (v. 4.2014, RSA*core*; LUMC, The Netherlands) [17], [18]. The migration was measured between baseline and follow-up, which in experiment C resulted in migration measurement between one and five years postoperative.

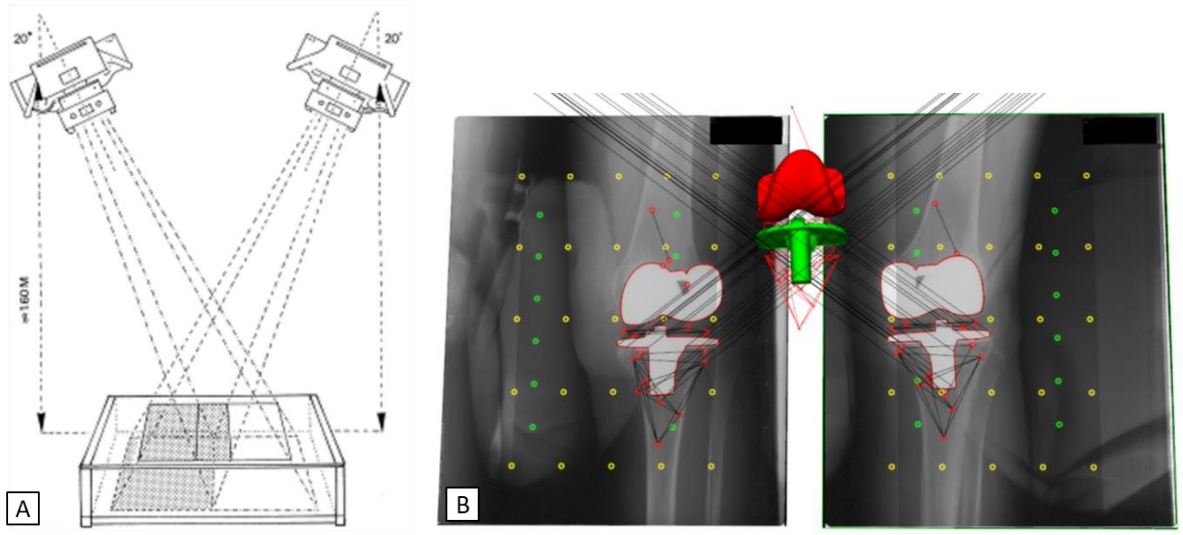


Figure 4: A) The RSA set-up consists of two separate roentgen tubes with corresponding digital radiography detectors and one calibration cage. The imaging body part is positioned on top of the calibration cage at the joint of interest of the X-ray bundles. Two separate roentgen images are taken virtually simultaneously. Image reproduced from [32]*. B) Screenshot of the model-based RSA software.*

3. RESULTS

3.1. Experiment A

3.1.1. Accuracy of V3MA and RSA

The accuracy ($t * RMS$) of V3MA varied between 0.02 mm and 0.15 mm for translations and was 0.03° for y-axis rotation (Table 2). The accuracy of RSA varied between 0.09 mm and 0.33 mm for translations and was 0.25° for y-axis rotation (Table 2). All data was normally distributed except for Ty in V3MA ($p < 0.01$) due to one outlier (Ty min: -0.12) (Appendix C). The differences between V3MA and RSA were normally distributed ($p = 0.98$).

For V3MA, the absolute mean differences were less than 0.03 mm for translations and 0.00° for rotations, which were practically not different from zero. For RSA, the mean differences were -0.01 mm (Tx), 0.02 mm (Ty), -0.11 mm (Tz), and 0.00° (Ry), indicating a systematic error (bias) in translations along the z-axis. Comparing the differences in accuracy between V3MA and RSA, V3MA showed a significantly better accuracy for Tz ($p = 0.01$). Furthermore, V3MA showed significantly less variance in the micromanipulator experiment for Tx ($p = 0.04$), Tz ($p = 0.02$), and Ry ($p < 0.01$) (Table 2).

Table 2: Accuracy of V3MA and RSA for imposed translations (n=5) along the x-,y- and z-axis and rotations around the y-axis (n=6)

Accuracy	V3MA				RSA			
	Translation (mm)			Rotation(°)*	Translation (mm)			Rotation(°)*
	Tx	Ty	Tz	Ry	Tx	Ty	Tz	Ry
t*RMS	0.05	0.15	0.02	0.03	0.09	0.09	0.33	0.25
Mean	0.01	-0.03	0.00†	0.00	-0.01	0.02	-0.11†	0.00
SD	0.01†	0.05	0.01†	0.01†	0.03†	0.03	0.05†	0.11†
Minimum	0.01	-0.12	-0.02	-0.01	-0.04	-0.01	-0.18	-0.14
Maximum	0.03	-0.01	0.00	0.02	0.04	0.05	-0.08	0.12
Shapiro-Wilk^	0.10	0.00^	0.99	0.52	0.39	0.71	0.93	0.67

**Micromanipulator only rotates around y-axis*
†Significant difference (paired t-test or Levene's test) between V3MA and RSA
^The test rejects the hypothesis of normality when the p-value is less than or equal to 0.05

3.1.2. Agreement between methods (V3MA and RSA) and micromanipulator

The limits of agreement between V3MA and micromanipulator did not exceed ±0.13 mm and ±0.06° for translations and rotations, respectively (Figure 5). There was no indication of a systematic or proportional bias. However, there seems to be a negative trend in Ty due to one outlier.

Furthermore, the limits of agreement between RSA and micromanipulator did not exceed ±0.18 mm and ±0.24° for translations and rotations, respectively (Figure 6). As for Tz, a significant systematic difference of -0.05 mm was observed, i.e., RSA consistently underestimated Tz (Figure 6). All limits of agreement between V3Ma and micromanipulator were smaller than those of RSA and micromanipulator, except for translations along the y-axis.

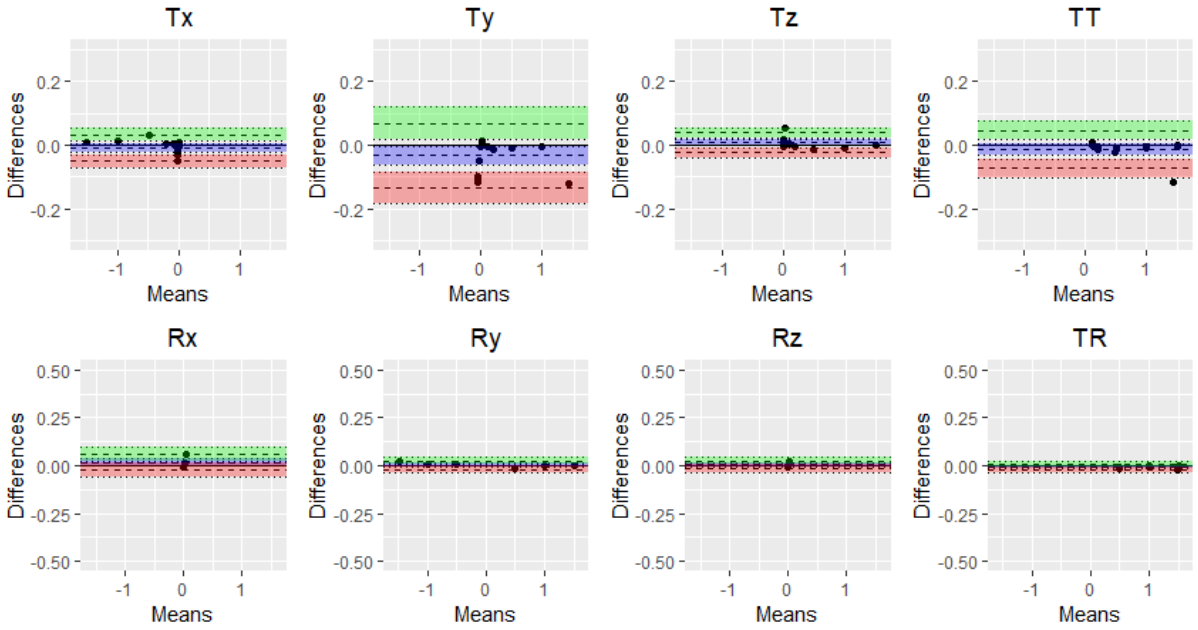


Figure 6: Bland-Altman plots of experiment A for comparison of V3MA and the micromanipulator. The limits of agreement (LOA, dashed lines) are presented as well as the 95%CI around the positive LOA (green area) and the negative LOA (red area).

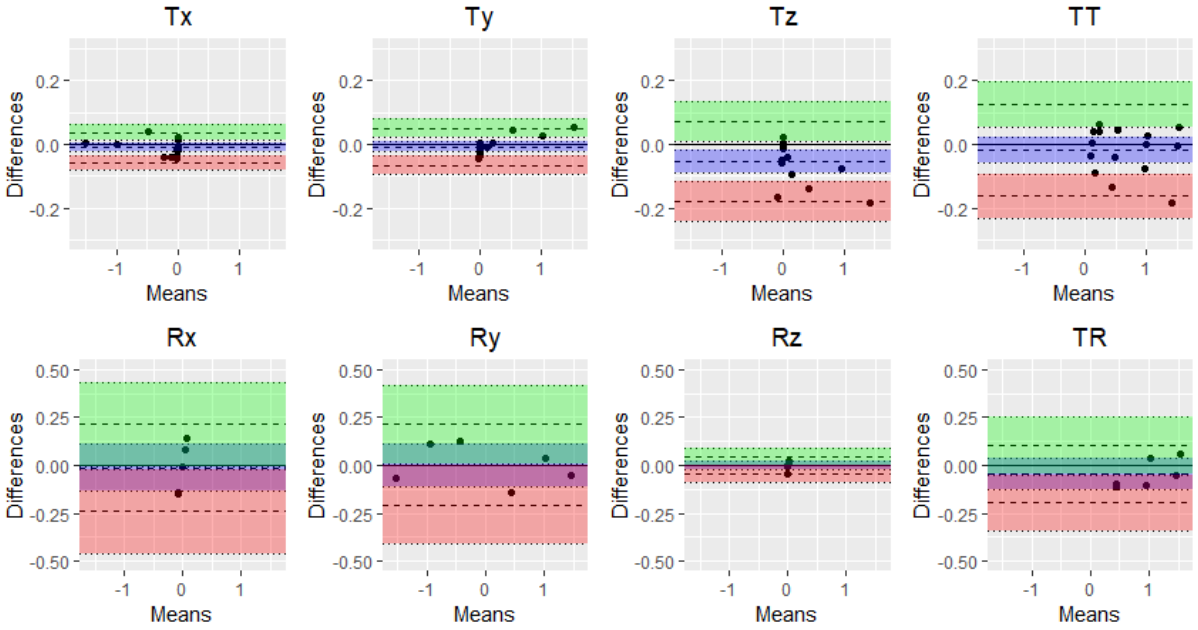


Figure 5: Bland-Altman plots of experiment A for comparison of RSA and the micromanipulator. The limits of agreement (LOA, dashed lines) are presented as well as the 95%CI around the positive LOA (green area) and the negative LOA (red area).

3.2. Experiment B

3.2.1. Precision of V3MA and RSA

V3MA precision (SD) ranged from 0.01 to 0.06 mm for translations and from 0.02 to 0.07° for rotations (Table 3). RSA precision (SD) ranged from 0.01to 0.06 mm for translations and between 0.04-0.25° for rotations (Table 4). Some of the data was not normally distributed: Tx and Tz in V3MA, and Tz and Rx in RSA (Table 4), as also shown in the density plots (Appendix C). The differences between V3MA and RSA were normal distributed except for Tz ($p = 0.04$) and Rx ($p = 0.03$).

When comparing the variance of V3MA and RSA, V3MA showed significant better precision (SD) in all translations and rotations, except for Ty: Tx ($p < 0.01$), Ty ($p < 0.01$), Tz ($p = 0.03$), Rx ($p < 0.01$), Ry ($p = 0.02$) , Rz ($p = 0.03$).

Table 3: Precision of V3MA for repeated measurements (n=16)

	Translations (mm)				Rotations (°)			
	Tx	Ty	Tz	TT*	Rx	Ry	Rz	TR*
95% CI (±2.131*SD)	-0.03 to 0.02	-0.14 to 0.11	-0.04 to 0.04		-0.03 to 0.05	-0.15 to 0.13	-0.03 to 0.03	
Mean	0.00	-0.02	0.00	0.05	0.01†	-0.01	0.00	0.05
SD	0.01†	0.06†	0.02†	0.04	0.02†	0.07†	0.02†	0.05
Minimum	-0.03	-0.14	-0.03	0.01	-0.02	-0.18	-0.03	0.01
Maximum	0.01	0.06	0.05	0.14	0.06	0.11	0.04	0.18
Shapiro-Wilk^	0.05^	0.07	0.01^	0.04^	0.11	0.21	0.92	0.00^

**TT and TR were calculated with 3D Pythagorean theorem*

†Significant difference with RSA (paired t-test or Levene's test)

^The test rejects the hypothesis of normality when the p-value is less than or equal to 0.05

Table 4: Precision of RSA for repeated measurements (n=12)

	Translations (mm)				Rotations (°)			
	Tx	Ty	Tz	TT*	Rx	Ry	Rz	TR*
95% CI (±2.110*SD)	-0.05 to 0.05	-0.03 to 0.03	-0.13 to 0.12		-0.13 to 0.29	-0.57 to 0.50	-0.09 to 0.07	
Mean	0.00	0.00	0.00	0.05	0.08†	-0.04	-0.01	0.20
SD	0.03†	0.01†	0.06†	0.04	0.10†	0.25†	0.04†	0.20
Minimum	-0.04	-0.02	-0.08	0.01	-0.02	-0.69	-0.09	0.03
Maximum	0.05	0.03	0.15	0.15	0.32	0.33	0.07	0.77
Shapiro-Wilk^	0.82	0.66	0.04^	0.03^	0.03^	0.10	0.68	0.00^

**TT and TR were calculated with 3D Pythagorean theorem*

†Significant difference with V3MA (paired t-test or Levene's test)

^The test rejects the hypothesis of normality when the p-value is less than or equal to 0.05

3.3. Experiment C

3.3.1. Clinical proof of concept

V3MA was feasible in all 6 patients regardless of the metal artifacts present in the CT images. Details of the CT images were provided (Appendix D). No significant bias was observed in the mean differences between V3MA (NL) and RSA (Figure 7). The limits of agreement did not exceed ±0.5 mm and ±0.7° for translations and rotations, respectively. The mean differences (limits of agreement) between V3MA (NL) and RSA were -0.05 mm (-0.49 to 0.39) and -0.14° (-0.79 to 0.51) for TT and TR, respectively.

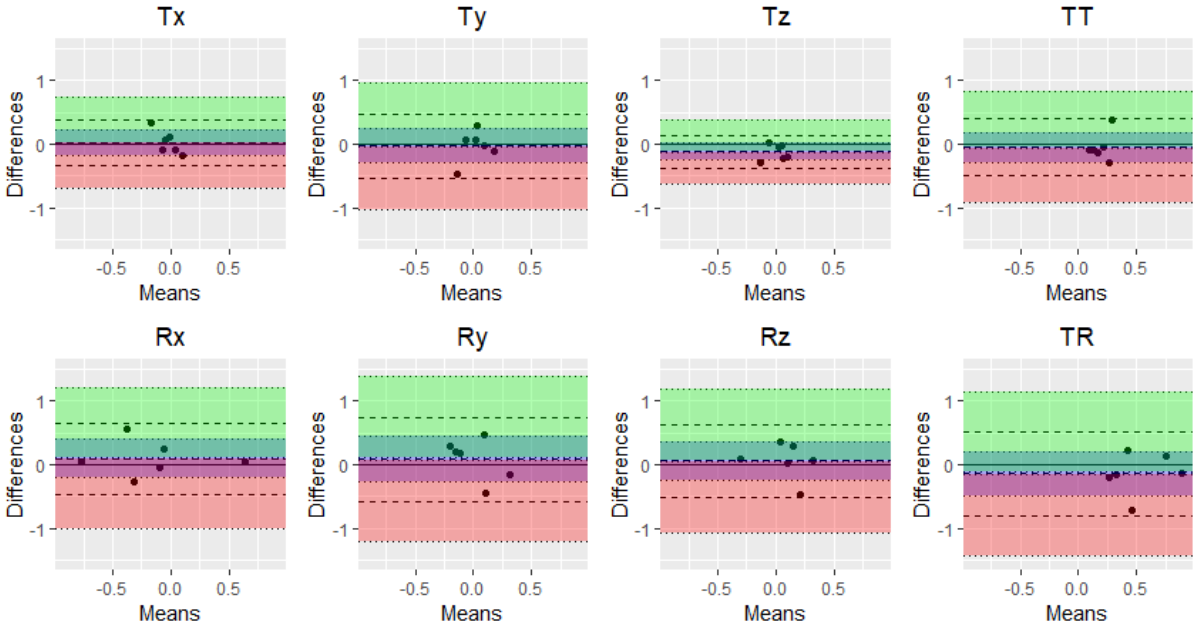


Figure 7: Bland-Altman plots of 6 patients for comparison between two methods: V3MA (NL) and RSA. The limits of agreement (LOA, dashed lines) are presented as well as the 95%CI around the positive LOA (green area) and the negative LOA (red area).

3.3.2. Exploratory interobserver variability

There was a good correlation between the two observers (ICC: 0.76, 95%CI: 0.58 to 0.87), indicative of a good interrater reliability. The limits of agreement between the two observers did not exceed ±0.3 mm and ±0.6° for translations and rotations, respectively (Figure 8). The mean differences did not differ significantly, suggesting that there was no systematic bias between the two observers. All limits of agreement between the two observers were smaller than the limits of agreement between V3MA and RSA. One outlier was observed in Rx (difference: 0.67°) without a clear explanation.

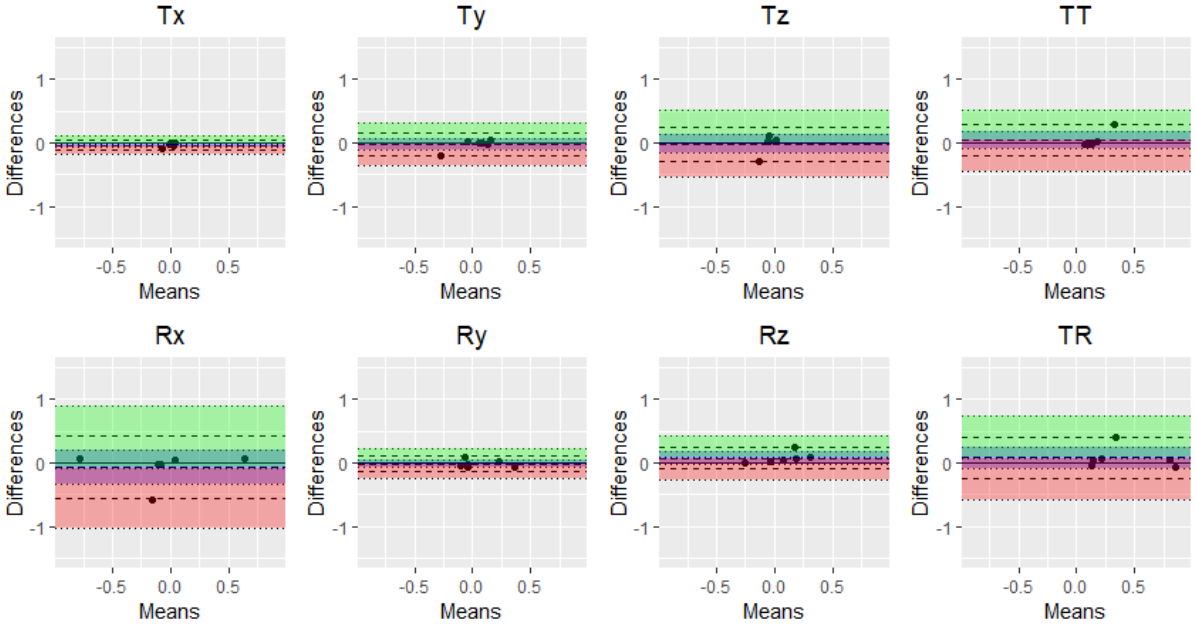


Figure 8: Bland-Altman plots of 6 patients for comparison between two observers: V3MA (NL) and V3MA (BK). The limits of agreement (LOA, dashed lines) are presented as well as the 95%CI around the positive LOA (green area) and the negative LOA (red area).

4. DISCUSSION

4.1. Key findings

We validated a novel, marker-free, CT-based, Volumetric Matching Migration Analysis (V3MA) in tibial components to measure migration, and compared its accuracy and precision to RSA.

In an experimental setting, the accuracy ($t * RMS$) of V3MA (Tx 0.05 mm; Ty 0.15 mm; Tz 0.02 mm; Ry 0.03°) was comparable to RSA (Tx 0.09 mm; Ty 0.09 mm; Tz 0.33 mm; Ry 0.25 mm). V3MA was more accurate for translations along the z-axis (Tz) and rotations around the y-axis (Ry), which could be explained by the true 3D imaging technique. It is known that out of plane motion (Tz, Ry, Rx) is less accurate in RSA [18]. RSA results were in line with the literature [17].

Furthermore, in another experiment, V3MA precision for translations (SD: 0.01 to 0.06 mm) and rotations (SD: 0.02 to 0.07°) was comparable to translations (SD: 0.01 to 0.05 mm) and rotations (SD: 0.03 to 0.21°) of RSA. There were significant differences between V3MA and RSA, but some of these differences can be interpreted as similar when taken the clinically relevance into account. The precision of V3MA was the lowest for translations along the y-axis (Ty). This slightly lower precision in the y-axis was expected, since the slice thickness is 0.5 mm, while the voxel size in x- and z-axis was 0.301 mm.

In a clinical setting, the mean differences between V3MA and RSA were close to zero and the limits of agreement did not exceed ±0.5 mm and ±0.7°. According to the literature, two migration measurement methods with limits of agreement which do not exceed ±0.5 mm and ±0.8° are considered equivalent [33]. Furthermore, the reliability between two observers using V3MA was good (ICC: 0.76, 95%CI: 0.58 to 0.87). This interrater reliability indicates that the method is robust for individual differences in the manual steps of the segmentation in V3MA. However, results should be interpreted with care since the clinical experiment included only six patients with good resolution CT images. Results may differ for other CT resolutions as suggested in the literature [21].

Two unexpected findings occurred: the -0.12 mm difference in accuracy measurement (Ty applied by micromanipulator: 1.5 mm and Ty measured by V3MA: 1.38 mm) and the difference in one datapoint between two observers in the interobserver variability (0.67° difference in Rx). The cause of both outliers was unknown, therefore they could not be removed from the results. Due to the first outlier, the differences of Ty in the micromanipulator experiment were not normally distributed. Testing for normality without the outlier resulted in a normal distribution of the differences between micromanipulator and measurement. The second outlier may be caused by the difference in order of registration, where BK used the align registration as starting position for the bone-bone registration.

Furthermore, in the precision experiment, one out of 4 set-ups had an unexpected different pixel size (0.98 mm instead of 0.31). This may have influenced precision of V3MA. Using a smaller pixel size, ought to result in a better registration and smaller precision values. This implicates that with consistent good pixel size, the precision of V3MA may be better.

In the results section, parametric descriptives (mean and SD) and parametric tests (t-test and Levene's test) were used to analyze the data despite some non-normal data. Parametric tests assume that the differences between pairs are normally distributed. However, tests for normality do not have much power in small sample sizes (rule of thumb: below 25) and visual tools to detect normality are not very helpful in detecting normality [34]. Also, normality may be assumed based on knowledge from other studies [25]. Therefore, we decided to use parametric descriptions and tests to compare the results of the current study to the literature.

4.2. Comparison of V3MA to the literature

The accuracy and precision of V3MA are consistent with previously published results on CTBMA [19], [21], [24], [25]. However, results are not directly comparable because other implant designs in other joints and other CTBMA methods were used. For example, Scheerlink et al. tested *CT-based spatial*

analysis (CTSA) in hip implants. They reported that the limits of agreement of their micromanipulator experiment for accuracy did not exceed ±0.28 mm and ±0.20° for translations and rotations, respectively [21]. These limits of agreement were in the same order as V3MA in knee implants, which did not exceed ±0.13 mm and ±0.06° for translations and rotations. Furthermore, Brodén et al. tested *CT-based Motion Analysis* (CTMA) in hip cups and reported that the accuracy ($t * RMS$) was in the range of 0.07-0.32 mm in translations and 0.21-0.82° for rotations [19]. The accuracy of CTMA in the humeral component of a shoulder implant was in the range of 0.07-0.23 mm in translations and 0.22-0.71° for rotations [25]. In comparison to CTMA, we found similar accuracy of V3MA in tibial implants, which ranged between 0.02-0.15 mm in translation and was 0.03° in y-axis rotation. Regarding the precision, Brodén et al. reported that the precision ($t * SD$) of CTMA on the humeral component was in the range of 0.08-0.15 mm in translations and 0.23-0.54° for rotation. The precision of CTMA on acetabular cups ranged between 0.01-0.09 mm in translations and 0.06-0.21° for rotation [19]. These results are in line with V3MA precision (SD): 0.01-0.06 mm in translations and 0.02-0.07° for rotation when multiplied by $t = 2.131$. The clinical precision of CTMA on acetabular cups was in the range of 0.07-0.31 mm in translation and 0.20-0.39° for rotation [24]. These precision results show that experimental precision is approximately 2 to 3 times better than clinical precision, as also reported previously [18]. We used CT images, but in theory, V3MA should also work with other 3D imaging techniques. One study showed that MRI was feasible for measuring tibial component migration [35]. It was, however, less precise than RSA.

4.3. Limitations

Although V3MA seems a promising alternative to RSA, the current study and the technique have some limitations. First of all, V3MA was only tested in one CT scanner using a specific CT protocol resulting in a certain CT resolution. As suggested before, the CT resolution could impact the accuracy and precision of V3MA.

Second, interobserver variability was determined between a master student in technical medicine (NL) with prior knowledge of segmentation and an expert in segmentation (BK). It remains to be seen if the interobserver variability is still as good when less experienced analysts use V3MA.

Third, V3MA accuracy was not determined for rotations around the x-axis (Rx) and the z-axis (Rz) because of the inability of the micromanipulator to rotate around these axes.

Fourth, we did not include the maximal total point motion (MTPM) in our measurements since it was not yet included in the current version of V3MA. MTPM represents the largest 3D migration of any point on the implant surface in mm. This value gives information about the magnitude of the migration and is used as a threshold value in RSA studies. An MTPM value of less than 0.5 mm at 6 months is considered acceptable, and more than 1.6 mm is considered unacceptable [36].

Furthermore, another limitation is that V3MA assumes rigid body displacement and therefore does not take possible changes in the shape or density of the bone into account. For example, over time, osteophytes, bone remodeling, stress shielding, and osteoporosis could change the aspect of the bone in follow-up imaging. Future long-term clinical studies should show if these bone changes interfere with the registration of V3MA.

4.4. Clinical implications

In the current experimental setting, V3MA outperformed RSA regarding accuracy and precision, except for Ty (superoinferior translation). Nevertheless, this direction was not inferior to Tz (anteroposterior translation), which was the least accurate and precise direction in RSA. Therefore, V3MA is a potential alternative tool for migration measurements of tibial components in TKA. For clinical use, other aspects should be considered as well. One advantage for V3MA is that CT scanners are widely available in contrast to RSA research set-ups, allowing for a more accessible method for migration measurement. Also, CT scans can be made in routine clinical practice. This increased

accessibility adds potential to evaluating new implants to improve patient safety. Another advantage of V3MA is that neither CAD models nor bone markers are necessary to measure migration. For example, one tibia 3D model was missing in experiment B which made migration measurement with model-based RSA difficult. In V3MA, this was not a problem since the CT images of the implant were used instead of a 3D model. On top of that, V3MA provides some additional functionality. For example, migration can be measured outside clinical studies and potentially provide more diagnostic information in an individual patient. One downside of V3MA compared to RSA is the higher radiation dose. Studies reported an effective dose of 0.16 mSv for a knee CT scan and 0.003 mSv for an RSA image of the knee [37], [38]. A total effective dose of 1 mSv (in case of 6 CT scans) is acceptable in studies when it "increases in knowledge leading to health benefit" [39]. Other aspects, such as patient satisfaction and costs can also be of interest.

4.5. Implant evaluation

There are several ways to evaluate the performance of TKA implants. The most relevant aspects used in clinical studies are patient satisfaction, durability, and safety. It follows that there is no clear definition of what constitutes a successful procedure.

Patient satisfaction: One of the main ways to evaluate the effectiveness of TKA is by assessing patient satisfaction. This patient perspective is measured by patient-reported outcome measures (PROMs). The 4 most commonly studied PROMs are the Oxford Knee Score, Knee Society Score, Osteoarthritis Outcome Score, and Western Ontario and McMaster Universities Osteoarthritis Index [40]. These questionnaires all broadly assess pain and activity function. Although the patient perspective remains important, PROMS are also susceptible to subjectivity and require large sample sizes with relatively large follow-ups.

Durability: It is also critical to evaluate how durable an orthopedic implant is. This durability can be assessed by looking at how long the implant lasts before it needs to be replaced. Registries, for instance, report the major revision percentages at 1, 3, 5, and 10 year(s). Consequently, it takes years to collect this data. In addition, RSA is used to measure migration and predict future loosening within two years of follow-up [14], [36]. However, migration measurement only accounts for loosening and does not consider infection risk or patient satisfaction. Therefore, the results of migration measurement studies are complementary to registries [8].

Safety: Another helpful way to evaluate the performance of implants is by looking at how safe it is. Items regarding patient safety include complications and clinical outcomes, which are also provided by registries. Additionally, RSA provides insight into early migration pattern of implants (within one to two years) and with relatively small patient cohorts (i.e., thirty to forty patients) [41]. As the turnover of new implants is high, such an early measurement could be beneficial before widespread use, thus contributing to patient safety.

The Orthopaedic Data Evaluation Panel (ODEP) combines multiple aspects of evaluating implant performance. This rating system combines registry data, peer-reviewed publications, and in-house manufacturer-initiated studies. The rating includes the number of years of use, evidence strength, and star's absence or presence. This star denotes a benchmark replacement rate: <5% at 10 years follow-up for TKA. In newly added implants where long-term evidence is absent, implant migration can be supplemental to ODEP.

4.6. Future research

Future research should include the clinical precision of V3MA. Unfortunately, we did not have double examinations to determine clinical precision. In addition, the influence of CT resolution on V3MA should be investigated, and MTPM should be added to the software. Furthermore, we are conducting a clinical study to investigate the limits of agreement between V3MA and RSA in a larger sample size.

5. CONCLUSION

In conclusion, V3MA has comparable accuracy and precision to standard model-based RSA and good interrater reliability. Therefore, this marker-free CT-based method could be used as an alternative to RSA to evaluate tibial component migration in TKA. However, further clinical studies with V3MA on more patients and different joints are necessary.

REFERENCES

[1] S. M. Kurtz *et al.*, "International survey of primary and revision total knee replacement," *Int. Orthop.*, vol. 35, no. 12, pp. 1783–1789, 2011, doi: 10.1007/s00264-011-1235-5.

[2] J. T. Evans, R. W. Walker, J. P. Evans, A. W. Blom, A. Sayers, and M. R. Whitehouse, "How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up," *Lancet*, vol. 393, no. 10172, pp. 655–663, 2019, doi: 10.1016/s0140-6736(18)32531-5.

[3] M. R. Norton, R. K. Vhadra, and A. J. Timperley, "The Johnson-Elloy (Accord) total knee replacement. Poor results at 8 to 12 years.," *J. Bone Joint Surg. Br.*, vol. 84, no. 6, pp. 852–5, 2002, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12211676.

[4] R. E. Gilbert, A. D. Carrothers, J. J. Gregory, and M. J. Oakley, "The St. Leger total knee replacement. A 10-year clinical and radiological assessment," *Knee*, vol. 16, no. 5, pp. 322–325, 2009, doi: 10.1016/j.knee.2009.02.005.

[5] NJR, "National Joint Registry for England Wales Northern Ireland and the Isle of Man. 18th Annual Report 2021," 2021.

[6] SAR, "The Swedish Arthroplasty Register Annual Report 2021," 2021.

[7] B. G. Pijls and R. G. H. H. Nelissen, "The era of phased introduction of new implants," *Bone Jt. Res.*, vol. 5, no. 6, pp. 215–217, 2016, doi: 10.1302/2046-3758.56.2000653.

[8] R. G. H. H. Nelissen, B. G. Pijls, and J. Karrholm, "RSA and Registries: The Quest for Phased Introduction of New Implants," *J. Bone Jt. Surg.*, vol. 3, pp. 62–65, 2011.

[9] H. Malchau, "On the importance of stepwise introduction of new hip implant technology : Assessment of total hip replacement using clinical evaluation, radiostereometry, digitized radiography and a national hip registry," Department of Orthopaedics, Institute of Surgical Sciences, Goteborg University, 1995.

[10] H. Malchau, "Introducing new technology: A stepwise algorithm," *Spine (Phila. Pa. 1976).*, vol. 25, no. 3, p. 285, 2000, doi: 10.1097/00007632-200002010-00004.

[11] M. Gross, "Innovations in surgery. A proposal for phased clinical trials," *J. Bone Jt. Surg. - Ser. B*, vol. 75, no. 3, pp. 351–354, May 1993, doi: 10.1302/0301-620x.75b3.8496198.

[12] J. Karrholm, B. Borssen, G. Lowenhielm, and F. Snorrason, "Does early micromotion of femoral stem prostheses matter? 4-7-year stereoradiographic follow-up of 84 cemented prostheses," *J. Bone Jt. Surg. - Ser. B*, vol. 76, no. 6, pp. 912–917, 1994, doi: 10.1302/0301-620x.76b6.7983118.

[13] B. G. Pijls *et al.*, "Early proximal migration of cups is associated with late revision in THA: A systematic review and meta-analysis of 26 RSA studies and 49 survival studies," *Acta Orthop.*, vol. 83, no. 6, pp. 583–591, 2012, doi: 10.3109/17453674.2012.745353.

[14] L. Ryd *et al.*, "Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses," *J. Bone Jt. Surg. - Ser. B*, vol. 77, no. 3, pp. 377–383, 1995, doi: 10.1302/0301-620x.77b3.7744919.

[15] E. R. Valstar, R. Gill, L. Ryd, G. Flivik, N. Börlin, and J. Kärrholm, "Guidelines for standardization of radiostereometry (RSA) of implants," *Acta Orthop.*, vol. 76, no. 4, pp. 563–572, 2005, doi: 10.1080/17453670510041574.

[16] G. Selvik, "Roentgen stereophotogrammetry: A method for the study of the kinematics of the skeletal system," *Acta Orthop.*, vol. 60, no. S232, pp. 1–51, 1989, doi: 10.3109/17453678909154184.

[17] B. L. Kaptein, E. R. Valstar, B. C. Stoel, P. M. Rozing, and J. H. C. Reiber, "A new model-based RSA method validated using CAD models and models from reversed engineering," *J. Biomech.*, vol. 36, no. 6, pp. 873–882, 2003, doi: 10.1016/S0021-9290(03)00002-2.

[18] B. L. Kaptein, E. R. Valstar, B. C. Stoel, H. C. Reiber, and R. G. Nelissen, "Clinical validation of model-based RSA for a total knee prosthesis," *Clin. Orthop. Relat. Res.*, no. 464, pp. 205–209, 2007, doi: 10.1097/BLO.0b013e3181571aa5.

[19] C. Brodén, H. Olivecrona, G. Q. Maguire, M. E. Noz, M. P. Zeleznik, and O. Sköldenberg, "Accuracy and Precision of Three-Dimensional Low Dose CT Compared to Standard RSA in Acetabular Cups: An Experimental Study," *Biomed Res. Int.*, vol. 2016, 2016, doi: 10.1155/2016/5909741.

[20] V. Otten, G. Q. Maguire, M. E. Noz, M. P. Zeleznik, K. G. Nilsson, and H. Olivecrona, "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.*, vol. 2017, 2017, doi: 10.1155/2017/3681458.
[21] T. Scheerlinck, M. Polfliet, R. Deklerck, G. Van Gompel, N. Buls, and J. Vandemeulebroucke, "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.*, vol. 87, no. 2, pp. 139–145, 2016, doi: 10.3109/17453674.2015.1123569.
[22] L. Olivecrona *et al.*, "Acetabular Component Migration In Total Hip Arthroplasty Using CT and A Semiautomated Program for Volume Merging," *Acta radiol.*, vol. 43, pp. 517–527, 2002, doi: 10.1034/j.1600-0455.2003.00086.x.
[23] C. Brodén, O. Sandberg, H. Olivecrona, R. Emery, and O. Sköldenberg, "Precision of CT-based micromotion analysis is comparable to radiostereometry for early migration measurements in cemented acetabular cups," *Acta Orthop.*, vol. 92, no. 4, pp. 419–423, 2021, doi: 10.1080/17453674.2021.1906082.
[24] C. Brodén *et al.*, "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.*, vol. 91, no. 3, pp. 260–265, 2020, doi: 10.1080/17453674.2020.1725345.
[25] C. Brodén *et al.*, "Accuracy and precision of a CT method for assessing migration in shoulder arthroplasty: an experimental study," *Acta radiol.*, vol. 61, no. 6, pp. 776–782, 2019, doi: 10.1177/0284185119882659.
[26] J. Vandemeulebroucke *et al.*, "Automated estimation of hip prosthesis migration: a feasibility study," in *Applications of Digital Image Processing XXXVI*, Sep. 2013, vol. 8856, p. 88561D, doi: 10.1117/12.2024382.
[27] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 196–205, Jan. 2010, doi: 10.1109/TMI.2009.2035616.
[28] O. Sköldenberg and M. Ödquist, "Measurement of migration of a humeral head resurfacing prosthesis using radiostereometry without implant marking: An experimental study," *Acta Orthop.*, vol. 82, no. 2, pp. 193–197, 2011, doi: 10.3109/17453674.2011.566133.
[29] J. M. Bland and D. G. Altman, "Comparing methods of measurement : why plotting difference against standard method is misleading," *Clin. Lab. Haematol.*, vol. 346, pp. 1085–1087, 1995.
[30] L. A. Koster, J. E. Meinardi, B. L. Kaptein, E. Van der Linden-Van der Zwaag, and R. G. H. H. Nelissen, "Two-year RSA migration results of symmetrical and asymmetrical tibial components in total knee arthroplasty: a randomized controlled trial," *Bone Joint J.*, vol. 103-B, no. 5, pp. 855–863, 2021, doi: 10.1302/0301-620X.103B5.BJJ-2020-1575.R2.
[31] N. N. de Laat, "CT-based migration analysis of tibial components in total knee arthroplasty.," University of Twente (unpublished master's thesis), 2022.
[32] E. R. Valstar, "Digital Roentgen Stereophotogrammetry: Development, Validation, and Clinical Application," 2000.
[33] K. T. van Hamersveld, P. J. Marang–van de Mheen, L. A. Koster, R. G. H. H. Nelissen, S. Toksvig-Larsen, and B. L. Kaptein, "Marker-based versus model-based radiostereometric analysis of total knee arthroplasty migration: a reanalysis with comparable mean outcomes despite distinct types of measurement error," *Acta Orthop.*, vol. 90, no. 4, pp. 366–372, Jul. 2019, doi: 10.1080/17453674.2019.1605692.
[34] R. Hoermann, J. E. M. Midgley, R. Larisch, and J. W. Dietrich, "Who is afraid of non-normal data? Choosing between parametric and non-parametric tests: A response," *Eur. J. Endocrinol.*, vol. 183, no. 2, pp. L1–L3, 2020, doi: 10.1530/EJE-20-0134.
[35] F. F. Schröder *et al.*, "Low-field magnetic resonance imaging offers potential for measuring tibial component migration," *J. Exp. Orthop.*, vol. 5, no. 1, 2018, doi: 10.1186/s40634-017-0116-2.
[36] B. G. Pijls, J. W. M. Plevier, and R. G. H. H. Nelissen, "RSA migration of total knee replacements: A systematic review and meta-analysis," *Acta Orthop.*, vol. 89, no. 3, pp. 320–328, 2018, doi: 10.1080/17453674.2018.1443635.
[37] D. Biswas, J. E. Bible, M. Bohan, A. K. Simpson, P. G. Whang, and J. N. Grauer, "Radiation exposure from musculoskeletal computerized tomographic scans," *J. Bone Jt. Surg. - Ser. A*, vol. 91, no. 8, pp. 1882–1889, 2009, doi: 10.2106/JBJS.H.01199.
[38] G. J. Teeuwisse W, Berting R, "Stralenbelasting bij orthopedische radiologie.," *Gamma*, vol. 8/9, pp. 197–200, 1998.

[39] European Commission, "Radiation protection 99. Guidance on medical exposures in medical and biomedical research," 1998.
[40] P. N. Ramkumar, J. D. Harris, and P. C. Noble, "Patient-reported outcome measures after total knee arthroplasty: A systematic review," *Bone Jt. Res.*, vol. 4, no. 7, pp. 120–127, 2015, doi: 10.1302/2046-3758.47.2000380.
[41] J. Kärrholm, R. H. S. Gill, and E. R. Valstar, "The history and future of radiostereometric analysis," *Clin. Orthop. Relat. Res.*, no. 448, pp. 10–21, 2006, doi: 10.1097/01.blo.0000224001.95141.fe.
[42] S. Klein and M. Staring, "The Elastix Manual," 2019. https://elastix.lumc.nl/doxygen/index.html (accessed Nov. 23, 2022).
[43] D. Eberly, "Euler angle formulas," *Geometric Tools*, 2020. http://www.geometrictools.com/Documentation/EulerAngles.pdf (accessed Dec. 19, 2022).

APPENDIX A: SCAN PROTOCOL (EXPERIMENT: A)

Bewaking vragen om mee te lopen naar Anatomie voor begeleiding (Route door kelders)...
Toestel 62888

16.45: ID 12345678g Anatomie 2485

Bij Anatomie, botten overladen naar blauwe box, vervoer naar Radiologie op karretje via de Kelders.

17.30:
Starten op CT

Metingen Fantoom (Micromanipulator)
X = Medial, Y = Anterior, Z = Lift-off.

Nummer / Tijd	As (x,y,z)	Translatie (mm)	X start	X eind	Y start	Y eind	Z start	Z eind
1 16.52	XYZ	0	0	0	0	0	0	0
2 18.06	Z	0.1	0	0	0.1	0	0	0.1
3 18.07	Z	0.2	0	0	0.2	0	0	0.2
4 18.08	Z	0.5	0	0	0.5	0	0	0.5
5 18.08	Z	1.0	0	0	1.0	0	0	1.0
6 18.09	Z	1.5	0	0	1.5	0	0	1.5
7 18.10	XYZ	0	0	0	0	0	0	0
8 18.11	rZ	-0.5	0	0	0	0	0	0
9 18.12	rZ	-1.0	0	0	0	0	0	0
10 18.13	rZ	-1.5	0	0	0	0	0	0
11 18.14	rZ	0	0	0	0	0	0	0
12 18.14	rZ	0.5	0	0	0	0	0	0
13 18.15	rZ	1	0	0	0	0	0	0
14 18.16	rZ	1.5	0	0	0	0	0	0
15 18.17	XYZ	0	0	0	0	0	0	0
16 18.17	Y	0.1	0	0,1	0	0	0,1	0
17 18.19	Y	0.2	0	0,2	0	0	0,2	0
18 18.19	Y	0.5	0	0,5	0	0	0,5	0
19 18.20	Y	1.0	0	1,0	0	0	1,0	0
20 18.21	Y	1.5	0	1,5	0	0	1,5	0
21 18.38	XYZ	0	0	0	0	0	0	0
22 18.39	X	0.1	0,1	0	0	0	0	0
23 18.39	X	0.2	0,2	0	0	0	0	0
24 18.40	X	0.5	0,5	0	0	0	0	0
25 18.41	X	1.0	1,0	0	0	0	0	0
26 18.41	X	1.5	1,5	0	0	0	0	0
27 18.42	XYZ	0	0	0	0	0	0	0

Figure 9: Scan protocol micromanipulator experiment with V3MA.

Nummer scan	As waarover wordt getransleerd of geroteerd	Translatie (mm)	Rotatie (°)
1	XYZ	0	0
2	Z	0.1	0
3	Z	0.2	0
4	Z	0.5	0
5	Z	1.0	0
6	Z	1.5	0
7	XYZ	0	0
8	Z	0	-0.5
9	Z	0	-1.0
10	Z	0	-1.5
11	Z	0	0
12	Z	0	0.5
13	Z	0	1.0
14	Z	0	1.5
15	XYZ	0	0
16	Y	0.1	0
17	Y	0.2	0
18	Y	0.5	0
19	Y	1.0	0
20	Y	1.5	0
21	XYZ	0	0
22	X	0.1	0
23	X	0.2	0
24	X	0.5	0
25	X	1.0	0
26	X	1.5	0
27	XYZ	0	0

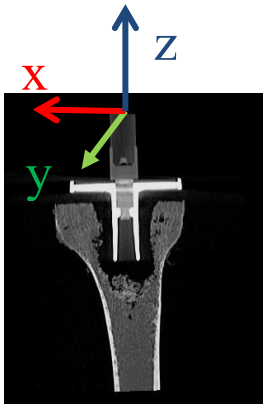
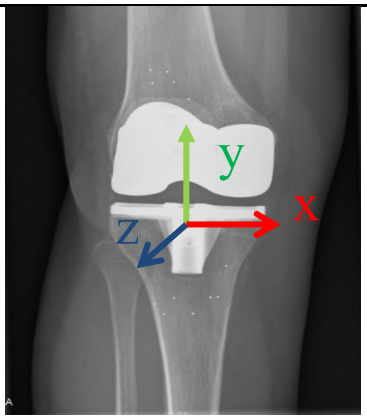
Figure 10: Scan protocol Micromanipulator experiment with RSA.

During the experiment the micromanipulator was first rotated left and then right in V3MA, in contrast to RSA where the micromanipulator first rotated right and then left. In de data the output was corrected for this input difference.

Coordinate system micromanipulator to V3MA (Table 5: Coordinate system of micromanipulator (left) and V3MA/RSA (right).Table 5):

- First, a translation to center of implant. Second, a rotation of -180° around the y-axis (right hand rule). Third, a rotation of -90° around the x-axis (right hand rule)

Table 5: Coordinate system of micromanipulator (left) and V3MA/RSA (right).

Micromanipulator	**V3MA / RSA**
	
Tz+	Ty+
Rz-	Ry+
Rz+	Ry-
Ty+	Tz+
Tx+	Tx-

APPENDIX B: TECHNICAL DETAILS V3MA

This section is an extract of N.N. de Laat's thesis for Technical Medicine (unpublished work) [31].

(1) Data acquisition

A three-dimensional gray-scale image $\boldsymbol{I}$ in spatial domain Ω assigns to every point $\boldsymbol{x} \in \Omega \subset \mathbb{R}^3$ a normalized gray value $\boldsymbol{I}(\boldsymbol{x}) \in [0,1]$ (called the intensity value at the point $\boldsymbol{x}$ of that image). Two images are needed for the registration: the moving image $\boldsymbol{I}_M = \boldsymbol{I}(\boldsymbol{x}) \forall\, \boldsymbol{x} \in \Omega_M \subset \mathbb{R}^3$ and the fixed image $\boldsymbol{I}_F = \boldsymbol{I}(\boldsymbol{x}) \forall\, \boldsymbol{x} \in \Omega_F \subset \mathbb{R}^3$. The fixed image is also called the baseline image ($t = t_0$) and the moving image is also called the follow-up image ($t = t_n$).

(2) Defining mask

A mask is defined in the fixed image (baseline) providing a labeled subset $\boldsymbol{L}$ of points $\boldsymbol{x} \in \boldsymbol{L}_F \in \Omega_F$. In V3MA three masks are used: align, prosthesis ($\boldsymbol{P}_F$), and bone ($\boldsymbol{B}_F$). *The align mask was not further specified in this appendix, because the align registration was only used to determine the starting position of bone and implant registration and not to calculate the migration*

(3) Registration

The underlying voxel intensities of $\boldsymbol{L}_F$ are used for the registration. Registration was performed using Elastix, which is an open source software for intensity-based medical image registration. The following components (with corresponding name in Elastix) were used in the multiresolution registration: interpolator (LinearInterpolator), metric (AdvancedNormalizedCorrelation), optimizer (AdaptiveStochasticGradientDescent), and transform (EulerTransform). In registration, the mapping from one image to the other results in a coordinate transformation $\boldsymbol{A}(\boldsymbol{x})$ that makes $\boldsymbol{I}_M(\boldsymbol{x})$ spatially aligned with $\boldsymbol{I}_F(\boldsymbol{x})$. In Elastix, for practical reasons, this transformation $\boldsymbol{A}$ was defined as a mapping from the fixed image to the moving image. The alignment optimization was estimated by minimizing the cost function (Equation (3)). [42]

$$\hat{\boldsymbol{A}} = arg \min_{\boldsymbol{A}} C(\boldsymbol{A}; \boldsymbol{I}_F, \boldsymbol{I}_M) \qquad (3)$$

(4) Calculation of migration

The migration of the tibial component relative to the tibia bone between baseline and follow-up image, was calculated by determining the relative movement of the two volumes (i.e., tibial component and tibia bone) with respect to a fixed migration coordinate system. The tibial component and tibia bone were assumed as rigid bodies, hence the transformations consisted of combinations of rotations and translations only. These transformations were expressed in homogenous coordinates to combine the translations and rotations in one matrix. The movement of a rigid body in three-dimensional space consisted of 6 degrees of freedom: 3 translations along the x-axis, y-axis, and z-axis and 3 rotations around these axes. These rotations were determined by the right hand rule and expressed in Euler angles, whilst the angles do not exceed 360°.

The following equations were used to calculate the migration:
The rigid transformation was defined as a transformation that, when acting on any vector $\boldsymbol{v}$, where $\boldsymbol{v} = [x, y, z, 1]^T$, and produces a transformed vector $\boldsymbol{v}'$ (Equation (4)).

$$\boldsymbol{v}' = \boldsymbol{A}\boldsymbol{v} \qquad (4)$$

$\boldsymbol{A}$ described in homogeneous coordinates gives a 4x4 matrix $\boldsymbol{A}^{[4,4]}$. The mathematical representation of the transformation of a point (x, y, z) towards another point (x', y', z') in a three-dimensional space is:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & p \\ r_{21} & r_{22} & r_{23} & q \\ r_{31} & r_{32} & r_{33} & r \\ 0 & 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

(5)

By definition, the estimated transformation matrix $\hat{\boldsymbol{A}}$ as calculated by Elastix is defined as a mapping from the fixed image $\boldsymbol{I}_F$ to the moving image $\boldsymbol{I}_M$, so that:

$$\boldsymbol{I}_M \sim \hat{\boldsymbol{A}}\, \boldsymbol{I}_F$$

(6)

Where $\hat{\boldsymbol{A}}$ is the estimated transformation resulting from alignment optimization. The transformation with a subset of pixels of the fixed image $\boldsymbol{L}_F$ for the alignment is given by:

$$\boldsymbol{L}_F{}' = \hat{\boldsymbol{A}}\, \boldsymbol{L}_F$$

(7)

Where $\boldsymbol{L}_F{}'$ is the transformed of $\boldsymbol{L}_F$ which Elastix tries to align to $\boldsymbol{L}_M$ so that:

$$\boldsymbol{L}_F{}' \sim \boldsymbol{L}_M$$

(8)

The reverse transformation yields:

$$\boldsymbol{L}_M{}' = \hat{\boldsymbol{A}}^{-1}\, \boldsymbol{L}_M$$

(9)

Where $\boldsymbol{L}_M$ represents a specific subset of the moving image which Elastix aligns to $\boldsymbol{L}_F$ (labeled by segmentation) to make $\boldsymbol{L}_M{}'$ (the transformed of $\boldsymbol{L}_M$) aligned with $\boldsymbol{L}_F$ (Figure 11).Figure 11 Elastix optimizes alignment so that:

$$\boldsymbol{L}_M{}' \sim \boldsymbol{L}_F$$

(10)

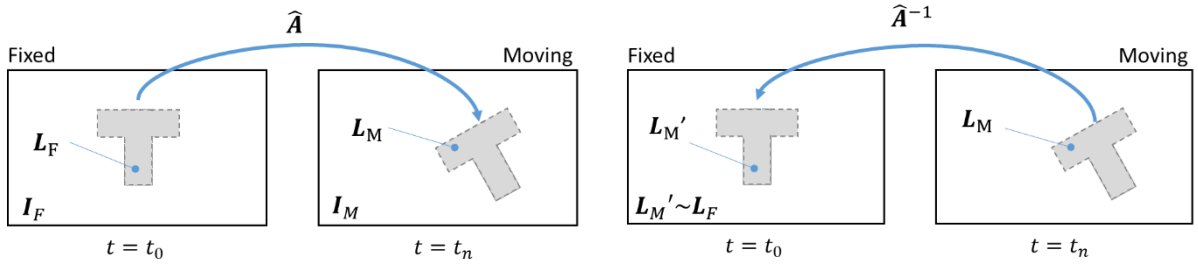


Figure 11: The estimated transformation $\hat{A}$ is defined as the mapping from the fixed image I_F to the moving image I_M (left image) and its reverse $\hat{\boldsymbol{A}}^{-1}$ from the moving to the fixed image (right image). The fixed image is also called the baseline ($t = t_0$) and the moving image is also called the follow-up ($t = t_n$). A specific subset of the fixed image L_F (labeled by segmentation) is used for the alignment optimization to a subset of the moving image L_M. Elastix optimizes alignment so that $L_M{}'$ (the transformed of L_M) is approximately equal to L_F.

In our study, Elastix optimized alignment of the bone in the moving image $\boldsymbol{B}_\mathrm{M}$ and the bone in the fixed image $\boldsymbol{B}_\mathrm{F}$ so that:

$$\boldsymbol{B}_M{}' \sim \boldsymbol{B}_F$$

(11)

Where $\boldsymbol{B}_\mathrm{M}{}'$ is the location where the bone ended up in the fixed image (Figure 12). In which the transformation yields:

$$\boldsymbol{B_M}' = \widehat{\boldsymbol{A}}_B^{-1}\, \boldsymbol{B_M} \tag{12}$$

Similarly, Elastix optimized alignment of the prosthesis in the moving image $\boldsymbol{P_M}$ and prosthesis in the fixed image $\boldsymbol{P_F}$ so that:

$$\boldsymbol{P_M}' \sim \boldsymbol{P_F} \tag{13}$$

Where $\boldsymbol{P_M}'$ is the location where the prosthesis ended up in the fixed image (Figure 12). Then the transformation is given by:

$$\boldsymbol{P_M}' = \widehat{\boldsymbol{A}}_P^{-1}\, \boldsymbol{P_M} \tag{14}$$

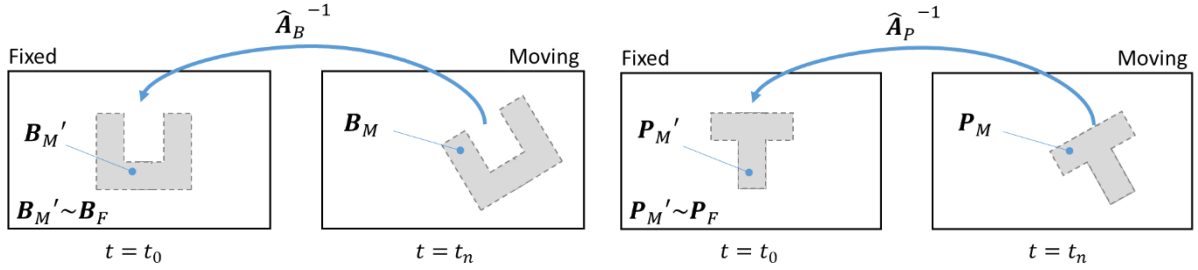


Figure 12: A graphical representation of the estimated reverse transformation of bone $\widehat{\boldsymbol{A}}_B^{-1}$ (left image) and prosthesis $\widehat{\boldsymbol{A}}_P^{-1}$ (right image) from the moving to the fixed image.

The relative migration over time of prosthesis with respect to the bone is of interest. Therefore, $\boldsymbol{P_M}$ (prosthesis in the moving image) is transformed with the bone transformation $\widehat{\boldsymbol{A}}_B^{-1}$ so that:

$$\boldsymbol{P_M}'' = \widehat{\boldsymbol{A}}_B^{-1}\, \boldsymbol{P_M} \tag{15}$$

Where $\boldsymbol{P_M}''$ is the transformed of $\boldsymbol{P_M}$ whilst using the reverse transformation of bone (Figure 13). Migration is then characterized by the residual alignment between $\boldsymbol{P_M}''$ and the initial transformed $\boldsymbol{P_M}'$ (which approaches $\boldsymbol{P_F}$, by optimization). The migration transformation $\widehat{\boldsymbol{A}}_{Mig}$ is given by:

$$\boldsymbol{P_M}'' = \widehat{\boldsymbol{A}}_{Mig}\, \boldsymbol{P_M}' \tag{16}$$

By substitution, the migration transformation $\widehat{\boldsymbol{A}}_{Mig}$ can hence be calculated:

$$\begin{aligned} \boldsymbol{P_M}'' &= \widehat{\boldsymbol{A}}_{Mig}\, \boldsymbol{P_M}' \\ \widehat{\boldsymbol{A}}_B^{-1}\, \boldsymbol{P_M} &= \widehat{\boldsymbol{A}}_{Mig}\, \widehat{\boldsymbol{A}}_P^{-1}\, \boldsymbol{P_M} \\ \widehat{\boldsymbol{A}}_B^{-1} &= \widehat{\boldsymbol{A}}_{Mig}\, \widehat{\boldsymbol{A}}_P^{-1} \\ \widehat{\boldsymbol{A}}_{Mig} &= \widehat{\boldsymbol{A}}_B^{-1}\, \widehat{\boldsymbol{A}}_P \end{aligned} \tag{17}$$

The transformation $\widehat{\boldsymbol{A}}_{Mig}$ is then transformed from the global coordinate system (provided by the CT image) to the migrating coordinate system (geometrical center of tibial component) (Equation (18)). The geometrical center is the calculated center of mass for a continuous mass distribution.

$$\widehat{\boldsymbol{A}}_{M0} = \boldsymbol{A}_0^{-1}\, \widehat{\boldsymbol{A}}_{Mig}\, \boldsymbol{A}_0 \tag{18}$$

Where $\boldsymbol{A}_0$ consists of a rotation of 90° around the x-axis and a translation to the center of the tibial component. The migration $\boldsymbol{M}$ is then derived from function f accepting matrix $\widehat{\boldsymbol{A}}_{M0}^{[4x4]}$ and returning $\boldsymbol{M}^{[6x1]}$ (Equation (19)). Migration consists of three translations parameters in millimeters (mm) and three rotation parameters in degrees (°): $\boldsymbol{M} = [T_x, T_y, T_z, R_x, R_y, R_z]$.

$$\boldsymbol{M} = f(\widehat{\boldsymbol{A}}_{M0}) \qquad (19)$$

This function f is implemented in V3MA software (*CTRSA_Migration_Calculation.py*, date: 14-04-2022, function: *matrix_to_parameters()*) and is based on Euler angle formulas [43]. The function finds the Euler angles from the rotation matrix and coverts these from radians to degrees.

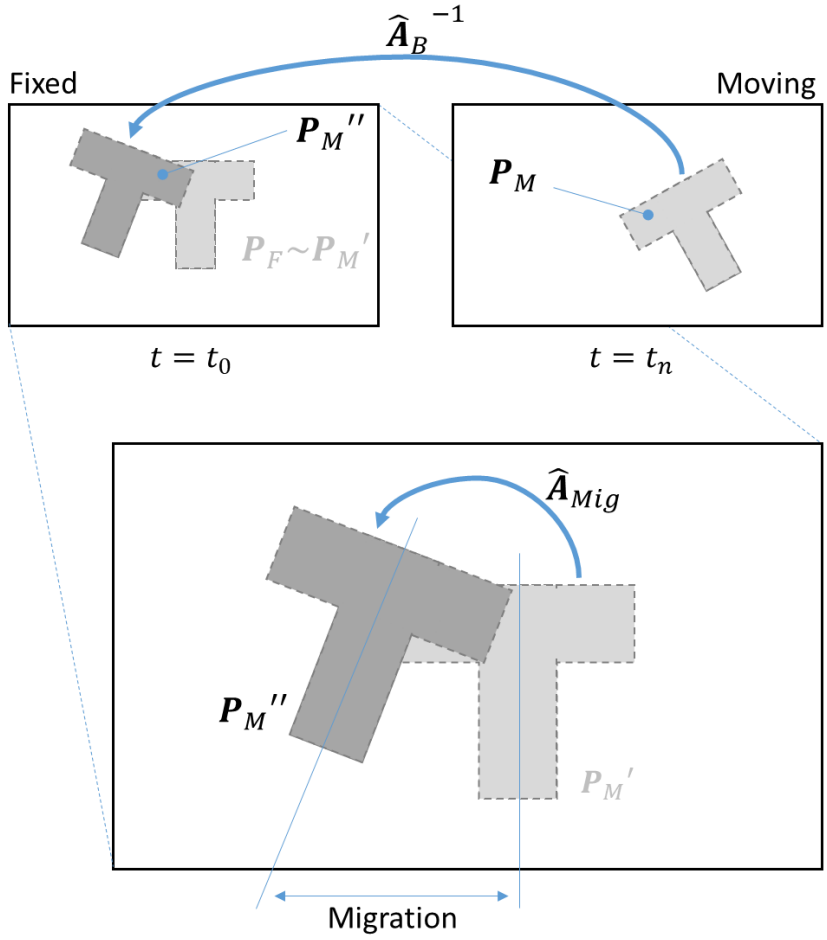


Figure 13: Migration is the relative displacement of prosthesis to the bone, between baseline and follow-up image.

APPENDIX C: DENSITY PLOTS (EXPERIMENTS: A, B)

Density plot experiment A

The mean difference between the micromanipulator and V3MA is close to zero. In comparison, the mean difference of the micromanipulator with RSA deviates from zero (Figure 14). Suggesting a better accuracy of V3MA. In addition, the variance of RSA is larger compared to V3MA. In V3MA one clear defined outlier is visible in the density plot.

Figure 14: Density plots accuracy (experiment A) of differences between measured migration and applied migration (n = 21) by the micromanipulator. The Colors correspond to the different axes. Delta = migration measurement – imposed migration.

Density plot experiment B – V3MA

Again we see density plots, but this time each axis is plotted separately. The variance between the repeated measurements seems greater in Set4 (Figure 15). In this set, the pixel size of the CT image was 0.98 mm in contrast to 0.31 mm of the CT images in the other 3 sets.

When comparing the density plots of V3MA (Figure 15) to the density plots of RSA (Figure 16), the variance of V3MA is smaller in all axes except in Ty. Suggesting a better precision in those axes.

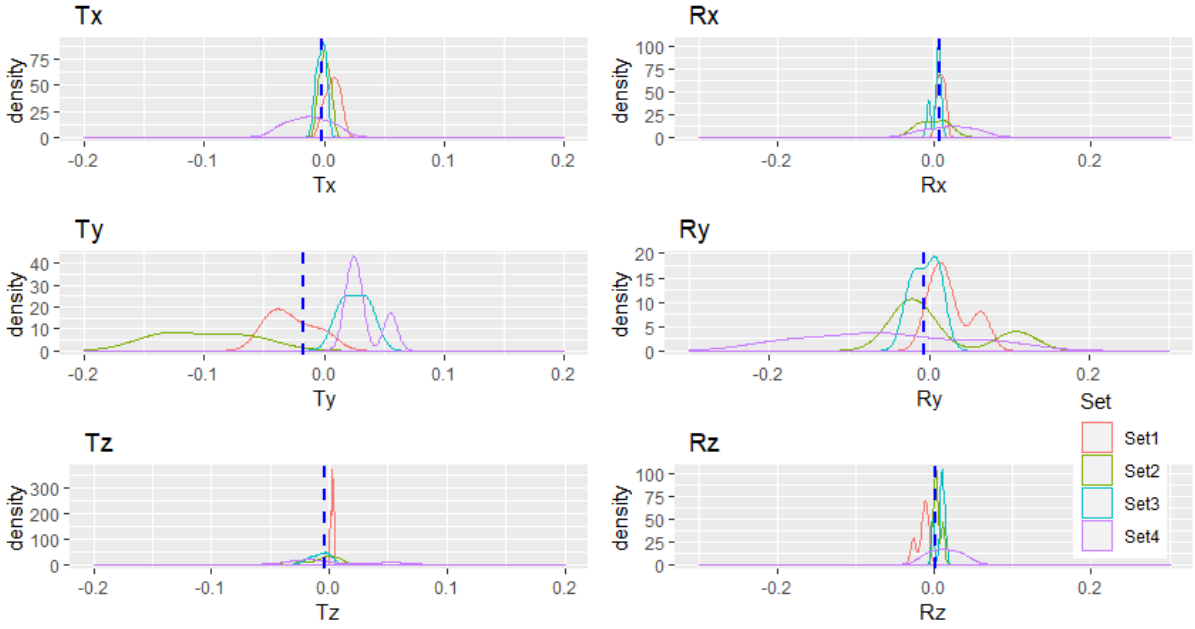


Figure 15: Density plots precision (experiment B) of V3MA (n=16). Colors correspond to specific cadaveric bone and tibial component combination (4 sets in total). The CT image of Set4 had different pixel size (0.98 mm instead of 0.31 mm).

Density plot experiment B – RSA

The combined mean difference of 3 sets deviated from zero in Rx, Ry, and Rz (Figure 16).

When comparing the density plots of RSA (Figure 16) to the density plots of V3MA (Figure 15), the variance of RSA is greater in all axes except in Ty. Suggesting a lesser precision in those axes.

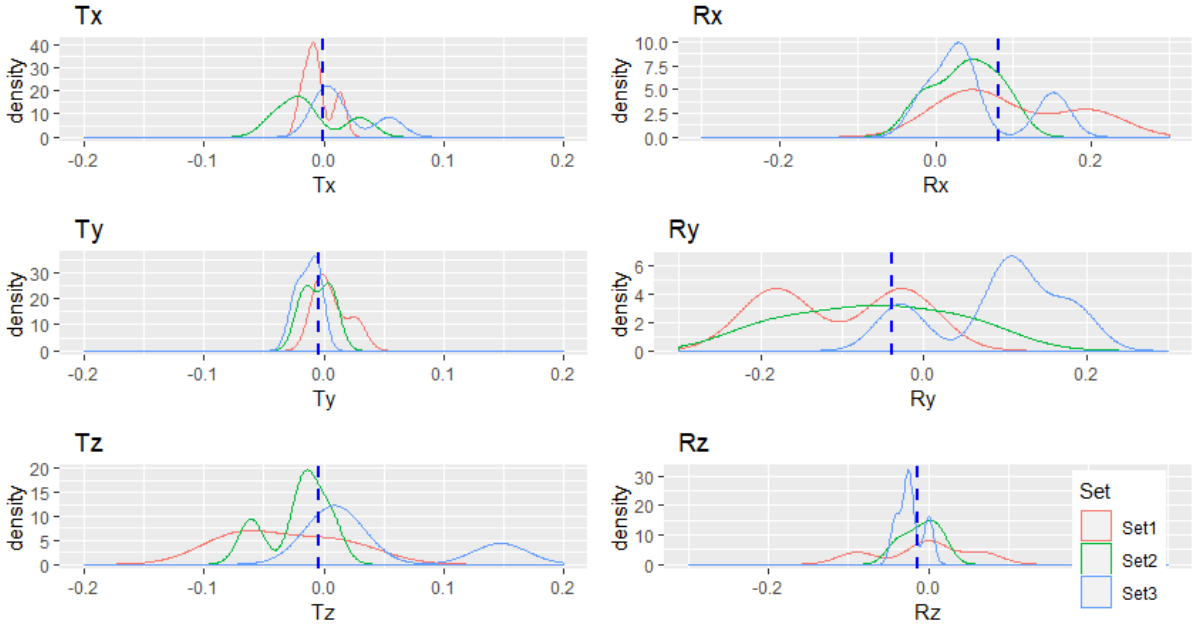


Figure 16: Density plots precision (experiment B) of RSA (n=12). Colors correspond to specific cadaveric bone and tibial component combination (3 sets in total).

APPENDIX D: CT PARAMETERS (EXPERIMENT C)

Table 6: Overview of CT parameters in experiment C. There are two CT images of each patient: one year (YR1) and five years (YR5) after primary total knee arthroplasty. All images have a slice thickness of 0.5 mm, matrix size 512x512 and convolution kernel FC30. Abbreviations: # = number of patient, PNS = patient number in study, FU = follow up, CT = Computed Tomography, kVp = kilo volt, mA = milli ampere, mm = millimeter, mGy = milli gray.

#	FU	PNS	CT scanner	Helical / volume	CT tube voltage (kVp)	Data collection diameter	CT tube current (mA)	Reconstruction diameter	Focal spot sizes (mm)	pixel spacing (mm)*	Dose (mGy)
1	YR1	22	Aquilion ONE	HELICAL_CT	135	320	80	190	0.9\0.8	0.37	7.2
	YR5		Aquilion ONE	HELICAL_CT	120	400	100	184.375	0.9\0.8	0.36	6.6
2	YR1	29	Aquilion ONE	VOLUME_CT	120	500	100	178	0.9\0.8	0.35	3.9
	YR5		Aquilion ONE	VOLUME_CT	120	400	90	182.812	0.9\0.8	0.36	4.6
3	YR1	45	Aquilion ONE	HELICAL_CT	120	500	110	191	0.9\0.8	0.37	7.2
	YR5		Aquilion ONE	VOLUME_CT	120	320	110	170.00	0.9\0.8	0.33	4.1
4	YR1	51	Aquilion ONE	VOLUME_CT	135	320	80	176	0.9\0.8	0.34	7.6
	YR5		Aquilion ONE	HELICAL_CT	120	320	110	193.75	0.9\0.8	0.38	4.1
5	YR1	61	Aquilion ONE	HELICAL_CT	120	400	80	185.937	0.9\0.8	0.36	5.8
	YR5		Aquilion ONE	HELICAL_CT	135	320	110	152.50	0.9\0.8	0.30	6.9
6	YR1	73	Aquilion ONE	VOLUME_CT	120	400	90	195.312	0.9\0.8	0.38	4.6
	YR5		Aquilion ONE	HELICAL_CT	135	500	66	158.203	0.9\0.8	0.31	3

APPENDIX E: RAW DATA (EXPERIMENTS: A, B, C)

Overview of raw data:

- Data experiment A: accuracy
- Data experiment B: precision
- Data experiment C: clinical proof of concept
- Multiple tables with limits of agreement

Table 7: Raw data experiment A – V3MA

V3MA micromanipulator								
#	Reference	Followup	Tx	Ty	Tz	Rx	Ry	Rz
1	000_NoAR	Tz0.1	-0.02	0.10	-0.01	-0.01	-0.01	0.00
2	000_NoAR	Tz0.2	-0.02	0.19	0.01	0.00	0.03	0.01
3	000_NoAR	Tz0.5	-0.03	0.49	0.01	-0.01	0.02	0.00
4	000_NoAR	Tz1.0	-0.02	0.99	0.02	-0.01	0.01	0.01
5	000_NoAR	Tz1.5	-0.05	1.38	0.05	0.01	0.01	-0.01
6	002_NoAR	r0.5	0.08	0.00	0.01	0.01	-0.49	0.00
7	002_NoAR	r1.0	0.17	0.00	0.01	0.02	-1.00	0.02
8	002_NoAR	r1.5	0.26	-0.12	0.05	0.06	-1.48	0.02
9	001_NoAR	r-0.5	-0.07	0.01	-0.02	0.00	0.49	0.00
10	001_NoAR	r-1.0	-0.13	0.01	-0.01	0.01	1.00	0.00
11	001_NoAR	r-1.5	-0.21	0.01	-0.02	0.02	1.50	0.00
12	003_NoAR	Ty0.1	0.01	0.00	0.11	0.00	0.00	0.00
13	003_NoAR	Ty0.2	-0.01	0.01	0.20	0.00	0.01	-0.01
14	003_NoAR	Ty0.5	-0.02	0.01	0.48	0.00	-0.01	-0.01
15	003_NoAR	Ty1.0	-0.03	0.01	0.99	0.00	0.00	-0.01
16	003_NoAR	Ty1.5	-0.02	0.02	1.50	0.00	0.01	-0.01
17	004_NoAR	Tx0.1	-0.09	-0.05	0.02	0.02	0.02	0.01
18	004_NoAR	Tx0.2	-0.19	0.00	0.00	0.00	0.01	0.00
19	004_NoAR	Tx0.5	-0.47	-0.10	0.02	0.03	0.01	0.01
20	004_NoAR	Tx1.0	-0.99	-0.11	0.01	0.03	0.02	0.01
21	004_NoAR	Tx1.5	-1.49	-0.12	0.00	0.03	0.03	0.01

Table 8: Raw data experiment A - RSA

RSA micromanipulator								
#	Reference	Followup	Tx	Ty	Tz	Rx	Ry	Rz
1	1	2	0.01	0.09	-0.05	-0.01	0.04	-0.02
2	1	3	-0.01	0.21	-0.16	0.05	-0.05	0.02
3	1	4	-0.02	0.55	0.02	0.17	-0.01	-0.02
4	1	5	-0.01	1.03	-0.01	0.05	0.09	-0.03
5	1	6	-0.01	1.55	0.01	0.10	-0.07	-0.02
6	7	8	0.05	-0.04	-0.02	-0.14	-0.37	0.02
7	7	9	0.16	-0.03	0.00	-0.15	-0.88	0.02
8	7	10	0.28	-0.02	-0.04	-0.01	-1.56	0.00
9	11	12	-0.06	0.01	-0.02	0.14	0.36	0.00
10	11	13	-0.16	-0.02	0.03	-0.01	1.04	-0.01
11	11	14	-0.21	0.00	0.01	0.09	1.45	-0.04
12	15	16	0.02	0.00	0.06	-0.15	-0.07	0.01
13	15	17	-0.02	-0.03	0.11	-0.17	0.01	0.02
14	15	18	-0.04	-0.03	0.36	-0.22	0.10	0.03
15	15	19	-0.04	-0.03	0.92	-0.15	0.02	0.04
16	15	20	-0.01	-0.03	1.32	-0.18	-0.03	0.03

17	21	22	-0.14	-0.01	-0.01	-0.08	0.12	0.00
18	21	23	-0.24	-0.02	-0.01	-0.07	0.07	0.02
19	21	24	-0.46	-0.01	-0.01	0.04	-0.16	-0.01
20	21	25	-1.00	-0.05	0.00	-0.08	-0.04	0.00
21	21	26	-1.49	-0.02	-0.06	0.00	-0.17	0.03

Table 9: Raw data experiment B - V3MA

V3MA precision									
#	Set	Reference	Followup	Tx	Ty	Tz	Rx	Ry	Rz
1	1	S01	S02	0.01	0.00	0.00	0.01	0.01	-0.01
2	1	S01	S03	0.00	-0.02	0.00	0.01	0.00	-0.01
3	1	S01	S04	0.01	-0.04	0.00	0.01	0.03	-0.01
4	1	S01	S05	0.01	-0.04	0.00	0.00	0.06	-0.03
5	2	S01	S02	0.00	-0.13	0.00	0.01	-0.01	0.01
6	2	S01	S03	0.00	-0.14	0.00	0.01	-0.03	0.00
7	2	S01	S04	0.01	-0.06	-0.03	-0.02	0.11	0.00
8	2	S01	S05	0.00	-0.08	-0.01	-0.01	-0.02	0.00
9	3	S01	S02	0.00	0.02	0.00	0.01	0.01	0.00
10	3	S01	S03	0.00	0.04	-0.02	-0.01	-0.01	0.01
11	3	S01	S04	0.00	0.01	0.00	0.01	0.01	0.01
12	3	S01	S05	-0.01	0.03	-0.01	0.00	-0.02	0.01
13	4	S01	S02	-0.03	0.02	-0.01	0.03	-0.04	0.01
14	4	S01	S03	0.01	0.03	-0.02	0.01	-0.18	-0.01
15	4	S01	S04	-0.01	0.06	-0.03	-0.02	0.08	0.02
16	4	S01	S05	-0.02	0.03	0.05	0.06	-0.09	0.04

Table 10: Raw data experiment B - RSA

RSA precision									
#	Set	Reference	Followup	Tx	Ty	Tz	Rx	Ry	Rz
1	1	1	2	-0.01	-0.01	-0.02	0.02	-0.03	0.01
2	1	1	3	-0.01	0.03	0.02	0.20	0.33	-0.09
3	1	1	4	0.01	0.00	-0.08	0.32	-0.69	0.07
4	1	1	5	-0.02	0.01	-0.07	0.07	-0.18	-0.01
5	2	1	2	-0.01	0.01	0.01	0.03	-0.10	0.00
6	2	1	3	-0.04	-0.02	-0.06	0.05	-0.03	-0.02
7	2	1	4	-0.02	-0.01	-0.01	0.09	0.06	0.02
8	2	1	5	0.03	0.00	-0.02	-0.02	-0.19	-0.04
9	3	1	2	0.00	-0.01	0.01	0.03	0.10	0.00
10	3	1	3	0.00	-0.02	0.15	0.15	0.11	-0.02
11	3	1	4	0.05	-0.02	0.00	0.03	0.18	-0.04
12	3	1	5	0.01	-0.01	0.02	-0.01	-0.03	-0.03

Table 11: Raw data experiment C: Tibial component migration between 1YR and 5YR (n=6).

Patient	V3MA								RSA							
	Translation (mm)				Rotation (°)				Translation (mm)				Rotation (°)			
	Tx	**Ty**	**Tz**	**TT***	**Rx**	**Ry**	**Rz**	**TR***	**Tx**	**Ty**	**Tz**	**TT***	**Rx**	**Ry**	**Rz**	**TR***
1	0.00	0.12	0.01	0.12	-0.09	-0.03	-0.03	0.10	-0.34	0.23	0.05	0.41	-0.66	-0.21	0.44	0.82
2	-0.12	-0.37	-0.27	0.48	-0.45	-0.05	0.30	0.54	-0.02	0.09	0.02	0.10	-0.18	-0.25	0.00	0.31
3	-0.02	0.18	0.00	0.19	-0.75	0.24	-0.26	0.82	-0.08	-0.11	0.20	0.24	-0.79	0.40	-0.35	0.95
4	0.00	-0.02	0.04	0.05	0.67	0.33	0.35	0.82	0.08	-0.10	0.06	0.14	0.62	-0.14	0.28	0.69
5	0.00	0.08	-0.05	0.10	0.06	-0.12	0.21	0.25	0.19	0.11	-0.06	0.23	-0.18	0.33	-0.14	0.41
6	0.04	0.06	-0.04	0.08	-0.12	-0.07	0.11	0.17	-0.07	-0.01	0.17	0.18	-0.07	-0.35	0.09	0.36

**TT and TR were calculated with 3D Pythagorean theorem*

Table 12: Detailed overview of the 4 sets used in Experiment B (in vitro precision).

	Set 1	Set 2	Set 3	Set 4
Femur	Femur 2509 Left: unknown	Femur 2485 Right: NG_Legacy_PS_Femur Sz F	Femur 2485 Right: LPS_FLEX_Femur Sz G	Femur 2509 Right: LPS_FLEX_Femur Sz E
Tibia	Tibia 2512 Right: LPS_FL_ST_Mobile Bearing Tibia Sz5	Tibia 2512 Left: LPS_FL_ST_Mobile Bearing Tibia Sz5	Tibia 2485 Left: LPS_FL_ST_Mobile Bearing Tibia Sz7	Tibia 2509 Left: unknown
Name in V3MA	*2509R2512R*	*2485R2512L*	*2485LeftTibia*	*2509LeftTibia*

Table 13: Limits of agreement of Bland-Altman plots in experiment A: V3MA & micromanipulator.

Bland-Altman - Experiment A: V3MA & mm (translations: n = 15) (rotations: n = 5)								
	Tx	Ty	Tz	TT	Rx	Ry	Rz	TR
Bias	-0.01	-0.03	0.01	-0.01	0.02	0.00	0.01	-0.01
Upper LOA	0.03	0.07	0.04	0.05	0.06	0.03	0.02	0.01
Lower LOA	-0.05	-0.13	-0.02	-0.07	-0.02	-0.02	-0.01	-0.02

Table 14: Limits of agreement of Bland-Altman plots in experiment A: RSA & micromanipulator.

Bland-Altman - Experiment A: RSA & mm (translations: n = 15) (rotations: n = 5)								
	Tx	Ty	Tz	TT	Rx	Ry	Rz	TR
Bias	-0.01	-0.01	-0.05	-0.02	-0.01	0.00	-0.00	-0.05
Upper LOA	0.04	0.05	0.07	0.12	0.21	0.21	0.04	0.11
Lower LOA	-0.06	-0.06	-0.18	-0.16	-0.24	-0.21	-0.05	-0.20

Table 15: Limits of agreement of Bland-Altman plots in experiment C: V3MA (NL) & RSA.

Bland-Altman - Experiment C: V3MA (NL) & RSA (n = 6)								
	Tx	Ty	Tz	TT	Rx	Ry	Rz	TR
Bias	0.02	-0.03	-0.12	-0.05	0.10	0.09	0.06	-0.14
Upper LOA	0.39	0.47	0.12	0.39	0.66	0.74	0.63	0.51
Lower LOA	-0.34	-0.53	-0.37	-0.49	-0.46	-0.57	-0.51	-0.79

Table 16: Limits of agreement of Bland-Altman plots in experiment C: V3MA (BK) & RSA

Bland-Altman - Experiment C: V3MA (BK) & RSA (n = 6)								
	Tx	Ty	Tz	TT	Rx	Ry	Rz	TR
Bias	0.06	-0.01	-0.11	-0.09	0.16	0.10	-0.02	-0.22
Upper LOA	0.39	0.34	0.14	0.14	0.65	0.79	0.49	0.28
Lower LOA	-0.28	-0.35	-0.35	-0.32	-0.31	-0.59	-0.53	-0.71

Table 17: Limits of agreement of Bland-Altman plots in experiment C: V3MA (NL) & V3MA (BK).

Bland-Altman - Experiment C: V3MA (NL) & V3MA (BK) (n = 6)								
	Tx	Ty	Tz	TT	Rx	Ry	Rz	TR
Bias	-0.03	-0.02	-0.01	0.04	-0.07	-0.01	0.08	0.08
Upper LOA	0.03	0.15	0.26	0.29	0.42	0.10	0.25	0.41
Lower LOA	-0.10	-0.19	-0.28	-0.20	-0.56	-0.13	-0.10	-0.25