Marker-based versus model-based radiostereometric analysis of total knee arthroplasty migration: a reanalysis with comparable mean outcomes despite distinct types of measurement error

Koen T VAN HAMERSVELD 1, Perla J MARANG–VAN DE MHEEN 2, Lennard A KOSTER 1, Rob G H H NELISSEN 1, Sören TOKSVIG-LARSEN 3, and Bart L KAPTEIN 1

1 Department of Orthopaedics, Leiden University Medical Center, Leiden, the Netherlands; 2 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands; 3 Department of Orthopaedics, Hässleholm Hospital, Hässleholm, Sweden and Department of Clinical Sciences, Lund University, Lund, Sweden
Correspondence: ktvanhamersveld@lumc.nl
Submitted 2018-10-30. Accepted 2019-03-25.

Due to the high accuracy and precision of radiostereometric analysis (RSA), late loosening of new implants can already be predicted with 2-year RSA results on small patient numbers (Ryd et al. 1995, Valstar et al. 2005, Nelissen et al. 2011).

RSA requires the bone and prosthesis to be accurately defined in 3 dimensions, usually achieved by inserting tantalum markers in the bone and by attaching or inserting markers (i.e., marker-based RSA). Prosthesis markers are generally inserted during surgery in the polyethylene of the implant (Kaptein et al. 2007). Alternatively, in model-based RSA the need for prosthesis markers is eliminated by matching a virtual projection of a 3D model with the contours of the radiographic projection of the implant (Kaptein et al. 2003). Results of model-based RSA are suggested to be comparable with conventional marker-based methods on a group level (Pijls et al. 2018), but direct comparisons on individual patient data are scarce (Kaptein et al. 2007, Hurschler et al. 2009).

We recently published the 2-year results of a randomized controlled trial (RCT) on implant migration of cemented metal-backed versus all-polyethylene tibial components in total knee arthroplasty (TKA) using the Triathlon TKA system.
Higher migration was found after 2 years for the metal-backed components, even though the difference was small. However, as migration measurements were based on markers inserted in the polyethylene, apparent migration of the modular metal-backed components may partly result from micromotion of the polyethylene insert with respect to the metal tray, a phenomenon that has been shown to occur in older fixed-bearing designs (Nilsson et al. 2003, Hansson et al. 2005).

In this study, we reanalyzed the metal-backed components with model-based RSA to eliminate any influence of modularity on migration results and thus investigate whether methodological differences between RSA methods would affect migration results. Second, we quantified movements of the polyethylene insert within the locking mechanism of the metal tray. Finally, we investigated whether the use of model-based RSA would result in different conclusions of the RCT as compared with the marker-based results.

**Patients and methods**

Full details of the original RCT regarding patients, randomization, follow-up, prosthesis, and surgical techniques have been described previously (van Hamersveld et al. 2018). Briefly, 2 surgeons implanted cemented, condylar-stabilizing, cruciate-retaining Triathlon total knee prostheses with either all-polyethylene ($n = 29$) or modular fixed-bearing metal-backed tibial components ($n = 30$). The metal tray was designed with a full peripheral capture locking mechanism and an anti-rotational central island (Łapaj et al. 2017). 2 patients with metal-backed components were analyzed with model-based RSA in the original RCT due to polyethylene marker occlusion, which precluded marker-based measurements. Hence, no marker-based results were available for comparison and these were thus excluded in the present study.

**Radiostereometric analysis**

The first RSA examination, performed on the first postoperative day, served as the reference for the migration measurements. Subsequent examinations were performed at 3 months, 1 year, and 2 years after surgery. RSA radiographs were performed in supine position with the knee in a biplanar calibration cage (cage 10, RSA Biomedical, Umeå, Sweden) and analyzed using Model-based RSA software version 4.1 (RSAcore, LUMC, Leiden, the Netherlands). For marker-based RSA analysis 5 tantalum markers (0.8 mm in diameter) were inserted during surgery, after drilling appropriate holes, at standardized positions in the polyethylene of both tibia designs. 2 markers were placed posteriorly, 2 anteromedially/anterolaterally, and 1 anteriorly. The number of markers available for migration calculations could differ over time due to marker occlusion (Figure 1a). Marker-based results of the metal-backed group were based on all 5 polyethylene markers in only 3 patients. As a result of marker occlusion in 1 or more follow-up moments, marker-based results were based on 4 polyethylene markers in 8 patients and on 3 markers in 17 patients. Model-based reanalysis was performed only in the metal-backed group, as the all-polyethylene components are radiolucent (Figure 1b). In the RSA analysis of the original report, a triangulated surface model (from reversed engineering, reduced to 5,000 triangles) was added for the tibial component and its virtual projections were matched with the contours of the radiographic projection of the implant. All other aspects of the analysis, such as insert markers, bone markers, and calibration markers, remained unchanged. Migration of the 28 metal-backed tibial components, by means of the 3D surface model, was calculated twice: with
the reference origin for migration calculations (1) in the geometric center of the model, which is the standard position for model-based RSA analysis, and (2) in the geometric center of the polyethylene markers, which is the standard position for marker-based RSA analysis (Figure 2). In addition, migration of the polyethylene insert markers was determined to assess whether the insert moved with respect to the metal tray. Lastly, method 2 allowed us to compare model-based metal-backed results with marker-based all-polyethylene results using the same reference origin. The precision of each RSA method was determined by means of double examinations at 1-year follow-up. The precision is expressed as the upper limits of the 95% confidence interval (CI) around zero motion (ISO 16087:2013(E) 2013). The primary outcome measure used in the original report is the maximum total point motion (MTPM), which is the length of the translation vector of the marker that moved the most. We also report the number of individual components showing “continuous migration,” defined by Ryd et al. (1995) as an increase in MTPM of ≥ 0.2 mm in the second postoperative year. The limits of marker stability (mean error) and scatter values (condition number) were set at 0.35 mm and 120, respectively, complying with the RSA guidelines (Valstar et al. 2005).

Statistics
We first estimated differences in model-based analyses with 2 different reference origins, i.e., the reference origin in the geometric center of the model versus the geometric center of the polyethylene markers, using regression analysis. Bland–Altman plots were constructed to determine the limits of agreement between the two RSA methods (Bland and Altman 1986). The limits of agreement, defined as the mean ± 1.96×SD, should be within ±0.5 mm of translation or ±0.8° of rotation for the measures to be considered equivalent. These thresholds were chosen as these are considered the smallest values of clinically relevant early migration when used as a predictor of aseptic loosening (Hurschler et al. 2009, Gudnason et al. 2017, Pijls et al. 2018). Boxplots were constructed to investigate micromotion of the polyethylene markers with respect to the metal tray along and about each orthogonal axis. Finally, an identical linear mixed-effects model as described in the original report (van Hamersveld et al. 2018) was used to analyze differences in migration between (model-based) metal-backed and (marker-based) all-polyethylene components while using the same reference origin (center of the polyethylene markers). As in the original report, log-transformation of outcome measures was applied when necessary to obtain normal distributions, and the same sensitivity analysis was performed given the unevenly distributed baseline characteristics sex and surgeon as possible confounders by adding these variables to the linear mixed-effects model (van Hamersveld et al. 2018). Significance was set at p < 0.05 (IBM SPSS Statistics 24.0; IBM Corp, Armonk, NY, USA).

Ethics, registration, funding, and potential conflicts of interest
The original study was approved by the Regional Ethical Review Board in Lund (entry no. 2013/434) and registered at isrctn.com (ID: ISRCTN04081530). All patients gave informed consent. The costs of the RSA radiographs made for the original study were supported by Stryker. The sponsor did not take part in the design, conduct, analysis, or interpretations stated in both the previous and current manuscript. The authors declare no competing interests.

Results
Double examinations were performed in 21 metal-backed components at 1-year follow-up to determine the precision of the RSA measurements. Model-based results were less precise in rotations, especially about the longitudinal axis (Table 1).

Marker-based versus model-based RSA
Regression analysis revealed that with (1) routine model-based RSA versus (2) model-based RSA with the reference origin in the geometric center of the polyethylene markers, the transverse, longitudinal, and sagittal translations were overestimated by 29% (CI 25–32), 7% (CI 0–13) and 26% (CI 24–28), respectively (illustrated for transverse translations in Figures 3a and 3b). As expected (for mathematical reasons, see Appendix), rotations and MTPM values were not influenced by the position of the reference origin and therefore identical between both model-based analyses. For fair com-
Comparison of marker-based and model-based translations, the reference origin for the model-based analysis was thus fixed at the geometric center of the polyethylene markers for the remaining analyses described below. This resolved the proportional bias (shown in Figure 3b and absent in Figure 3d) (Ludbrook 1997).

Comparing marker-based with model-based RSA, translations showed small limits of agreement indicating that both methods can be used interchangeably (Table 2). The limits of agreement for the rotations and MTPM were larger, especially for rotations about the longitudinal axis (Table 2).

**Micromotion of the polyethylene insert with respect to the metal tray**

Boxplots were constructed to investigate micromotion of the polyethylene insert with respect to the metal tray along and about each orthogonal axis at 3, 12, and 24 months’ follow-up (Figure 4). The majority of the measurements were within the 95% confidence interval of zero motion (i.e., the precision of the RSA method, indicated by the shaded areas in Figure 4) and group median values did not appear to increase over time. A few outliers depicted in Figure 4 were evaluated to determine the nature of the extreme values, all of which were...
found to be due to measurement error as a result of instability or occlusion of the polyethylene markers. The error of patient 6 was due to one polyethylene marker moving posteriorly close to the periphery of the drilled hole where it was inserted (resulting in a mean error between 0.31 and 0.33 at 3, 12, and 24 months, close to the limit of 0.35). This marker stabilized within 3 months, as the polyethylene micromotion values were close to zero when 3 months’ follow-up was taken as the reference (mean error between 0.02 and 0.03 at 12 and 24 months).

A similar cause was found in the analysis of patient 58, but in this case 2 anterior markers moved anteriorly; results were also close to zero when 3 months’ follow-up was taken as the reference. In the analysis of patient 22, patient 32, and patient 40, only 3 polyethylene markers were available of which 1 was partly occluded in 1 or more follow-up moments by either the tibial component or by another marker; slightly adjusting the position of these markers resulted in values close to zero in all directions.

**Change in results of original trial**

When repeating the analysis of the primary outcome (MTPM after 2 years of follow-up) of the original report (van Hamersveld et al. 2018) with the model-based migration values, comparable group differences were found: the all-polyethylene group had an MTPM (CI) of 0.61 (0.49–0.74) versus 0.81 (0.68–0.95) for the marker-based metal-backed group; and versus 0.82 (0.68–0.96) for the model-based metal-backed group (Figure 5).

In the original paper, continuous migration of ≥ 0.2 mm in MTPM in the second postoperative year was seen in 4 components in both groups. These 4 individual components of the metal-backed group showed similar migration patterns using model-based analysis (i.e., continuous migration in the second postoperative year). However, 2 additional metal-backed components showed continuous migration based on the model-based analysis. In both cases, the increase in MTPM in the second postoperative year was likely the result of a sudden increase in rotation about the longitudinal axis due to model-fit inaccuracies, as all other parameters remained stable (data not shown).

The other RSA parameters showed comparable between-group results when repeating the analysis with model-based migration values, except for translations along and rotations about the longitudinal (y-)axis, again, due to model-fit inaccuracies (Table 3, see Supplementary data).

In this trial, 2 surgeons performed the surgeries. When stratifying the results by surgeon as performed in the post hoc sensitivity analysis of the original report, the observed difference in MTPM in favor of the all-polyethylene design was smaller and not statistically significant. Repeating this sensitivity analysis with the model-based measurements resulted in similar conclusions (Table 4, see Supplementary data).
Discussion

We investigated whether model-based RSA, utilizing a different reference origin as compared with marker-based RSA, would affect migration outcomes. By doing so, we were also able to quantify movements of the polyethylene insert within the locking mechanism of the Triathlon metal tray and explore the disadvantages of each RSA method. If the results differed systematically, pooling and comparing RSA data from studies using different RSA techniques would be impaired unless adjusted for the methods being used. However, if the insert moves with respect to the metal tray in modular TKA, then marker-based migration values of the tibial component in the transverse plane are unreliable (Nilsson et al. 2003), and likely produce random error that cannot be corrected for when comparing with model-based RSA studies. Now that an increasing number of RSA studies are available with long-term follow-up, meta-analysis becomes possible—but one must ascertain pooling of data is justified when different RSA methods have been used.

Our study demonstrated systematic differences in translations but not rotations between model-based RSA and marker-based RSA. These differences are caused by the difference in reference origin that is used for migration calculation (Hurschler et al. 2009). As compared with the tibia 3D surface model, the origin in the center of the polyethylene markers overestimated the model-based transverse, longitudinal, and sagittal translations of the tibial component by 29%, 7%, and 26%, respectively. Correcting for this proportional bias, by using a factor or by using the same reference coordinate system in both analysis methods, resulted in nearly identical translations between model-based and marker-based analysis. For the rotations and MTPM values, the limits of agreement between marker-based and model-based RSA were larger because of the reduced precision of model-based rotations, particularly about the longitudinal axis. This is known and due to the relatively round, symmetrical shape of the tibial component in the transverse plane (Kaptein et al. 2007). Still, the limits of agreement between methods were within ± 0.5 mm and ± 0.8° and conclusions on the primary outcome of the RCT regarding group differences in MTPM remained unchanged. Furthermore, we found no evidence for the presence of insert micromotion and excluded this as a cause of unreliable marker-based migration measurements for the modular Triathlon TKA system. For the individual patient, however, use of a different method may result in substantial differences due to various types of measurement error (e.g., marker occlusion and model-fit inaccuracies). Therefore, one must not put too much weight on strict migration thresholds in individual patients (e.g., 0.2 mm of MTPM migration in the second postoperative year).

Our findings are in line with an earlier comparison between marker-based and model-based RSA (Hurschler et al. 2009). However, in that study, among other methodological differences, a uniplanar RSA setup was used resulting in marked differences in accuracy between “in-plane” and “out-of-plane” translations and rotations. In the present study we used a biplanar technique, and we did not find such a dichotomy in accuracy. Nevertheless, our findings further support their conclusion that model-based RSA can be used interchangeably with marker-based RSA, at least for the Triathlon TKA, provided that the same reference origin is used or corrected for using a factor when analyzing translations.

Previous studies evaluating insert micromotion relative to the metal tray in modular TKAs found small movements in Nuffield fixed-bearing TKAs (Corin Medical Ltd., UK) (Hansson et al. 2005) and NexGen fixed-bearing TKAs (Zimmer, USA) (Nilsson et al. 2003). In the latter study, these movements were closely examined and found to be greater in the transverse plane, which corresponds to the polyethylene–metal tray interface (Nilsson et al. 2003). This contrasts with our results and may be explained by the different designs of the locking mechanisms that were used. In a recent retrieval study of Łapaj et al. (2017), backside damage as a result of abrasion following micromotion of the polyethylene was found in designs with dovetail locking mechanisms, especially in the NexGen trays. Contrarily, they found no evidence for abrasion in the Triathlon knees owing to the full peripheral capture locking mechanism. Furthermore, the anti-rotational central island of the Triathlon design has been shown to effectively reduce micromotion to a minimum for a given reacted torque as compared with other TKA designs, including NexGen (Bhimji et al. 2010), although this mechanical study was performed by the research and development department of Stryker. It should be noted, however, that random error as a result of the reduced precision of model-based RSA limits firm conclusions on the presence of (longitudinal) rotations of the polyethylene within the locking mechanism. Nevertheless, the found translations were minimal and all outliers were found to be caused by polyethylene marker instability or occlusion, thus unlikely to be the result of micromotion in the polyethylene–metal tray interface.

A limitation of this study is that we compared the results of only one tibial component design. As the precision of model-based RSA depends on the shape and accuracy of the fitted model (Kaptein et al. 2003), differences between marker-based and model-based RSA results may be smaller or larger depending on the TKA design and also depending on the location of the prosthesis markers, either in the insert, or attached to the metal tibial component.

In summary, systematic differences in translations between marker-based and model-based RSA analysis disappeared when adjusted for the different reference origins being used for migration calculations. Micromotions of the polyethylene insert within the Triathlon metal tray were at most negligibly small. Mean migration results of model-based and marker-based measurements were comparable between groups when using the same reference origin, even though migration pat-
terns of individual patients may differ between RSA methods due to various types of measurement error.

**Supplementary data**

Tables 3–4 and the Appendix are available as supplementary data in the online version of this article, http://dx.doi.org/10.1080/17453674.2019.1605692

The study was designed and coordinated by KH and STL. Data collection was performed by KH. Statistical analysis was done by KH, PM, and BK. KH, PM, LK, RN, STL, and BK interpreted the data and wrote the initial draft manuscript. All authors critically revised and approved the manuscript.

The authors are grateful to Håkan Leijon for providing the marker-based RSA measurements and for his valuable help in performing the additional model-based RSA measurements.

*Acta* thanks Raed Itayem, Maiken Stilling and Lars Weidenhielm for help with peer review of this study.

Bhimji S, Wang A, Schmalzried T. Tibial insert micromotion of various total knee arthroplasty devices. J Knee Surg 2010; 23(3): 153-62.

Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476): 307-10.

Budnason A, Adalberth G, Nilsson K G, Hailer N P. Tibial component rotation around the transverse axis measured by radiostereometry predicts aseptic loosening better than maximal total point motion. Acta Orthop 2017; 88(3): 282-7.

Hansson U, Toksvig-Larsen S, Jorn L P, Ryd L. Mobile vs. fixed meniscal bearing in total knee replacement: a randomised radiostereometric study. The Knee 2005; 12(6): 414-18.

Hurschler C, Seehaus F, Emmerich J, Kaptein B L, 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.

ISO 16087:2013(E). Implants for surgery: Roentgen stereophotogrammetric analysis for the assessment of migration of orthopaedic implants. Geneva: International Organization for Standardization; 2013.

Kaptein B L, Valstar E R, Stoel B C, Rozing P M, Reiber J H. A new model-based RSA method validated using CAD models and models from reversed engineering. J Biomech 2003; 36(6): 873-82.

Kaptein B L, Valstar E R, Stoel B C, Reiber H C, Nelissen R G. Clinical validation of model-based RSA for a total knee prosthesis. Clin Orthop Relat Res 2007; 464: 205-9.

Łapaj Ł, Mróz A, Kokoszka P, Markuszewski J, Wendland J, Helak-Lapaj C, et al. Peripheral snap-fit locking mechanisms and smooth surface finish of tibial trays reduce backside wear in fixed-bearing total knee arthroplasty. Acta Orthop 2017; 88(1): 62-9.

Ludbrook J. Comparing methods of measurements. Clin Exp Pharmacol Physiol 1997; 24(2): 193-203.

Nelissen R G, Pijs B G, Karrholm J, Malchau H, Nieuwenhuijse M J, Valstar E R. RSA and registries: the quest for phased introduction of new implants. J Bone Joint Surg Am 2011; 93(Suppl. 3): 62-5.

Nilsson K G, Henrikson A, Dalen T. In vivo determination of modular tibial insert micromotion. Trans Orthop Res Soc 2003; 28: 1402.

Pijs B G, Plevier J W M, Nelissen R. RSA migration of total knee replacements. Acta Orthop 2018; 89(3): 320-8.

Ryd L, Albrektsson B E, Carlsson L, Dansgard F, Herberts P, Lindstrand A, et al. Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses. J Bone Joint Surg Br 1995; 77(3): 377-83.

Valstar E R, Gill R, Ryd L, Flivik G, Borlin N, Karrholm J. Guidelines for standardization of radiostereometry (RSA) of implants. Acta Orthop 2005; 76(4): 563-72.

van Hamersveld K T, Marang-van de Mheen P J, Nelissen R G H H, Toksvig-Larsen S. Migration of all-polyethylene compared with metal-backed tibial components in cemented total knee arthroplasty: a randomized controlled trial. Acta Orthop 2018; 89(4): 412-7.