Arthroplasty Today 6 (2020) 835–844

Contents lists available at ScienceDirect

Arthroplasty Today

journal homepage: http://www.arthroplastytoday.org/

Original Research

Clinical and Radiographic Outcome of Gap Balancing Versus Measured Resection Techniques in Total Knee Arthroplasty

Eva De Wachter, MD [a,*], Johan Vanlauwe, PhD, MD [a], Robert Krause, MD [b], Hans Bayer-Helms, MD [c], Dirk Ganzer, MD [d], Thierry Scheerlinck, PhD, MD [a]

[a] Department of Orthopedic Surgery and Traumatology, UZ Brussel, Brussels, Belgium
[b] Department of Orthopedic Surgery and Traumatology, Oberlinklinik, Potsdam, Germany
[c] Department of Orthopedic Surgery and Traumatology, Hilden, Germany
[d] Department of Orthopedic Surgery and Traumatology, Dietrich-Bonhoeffer-Klinikum, Altentreptow, Germany

ARTICLE INFO

Article history:
Received 26 March 2020
Received in revised form
29 July 2020
Accepted 30 July 2020
Available online xxx

Keywords:
TKA
Arthroplasty
Measured resection
Gap balancing
Survival
Knee Society Score

ABSTRACT

Background: There is no consensus regarding superiority between gap balancing (GB) and measured resection (MR) techniques to implant total knee arthroplasties. In a multicenter setup, we compared both techniques using the same prosthesis.
Methods: We included 262 balanSys posterior-stabilized total knee arthroplasties from 4 centers: 3 using the MR (n = 162) and one using the GB technique (n = 100), without navigation.
Results: There was no significant difference in the Knee Society Score or visual analog scale pain at 2- and 7-year follow-up. The visual analog scale for satisfaction was significantly better in the MR group at 2 but not at 7 years. We found a significantly higher average valgus in the GB group, but the overall alignment was within 2° of neutral on the full-leg radiographs. There were no significant differences concerning radiolucency and survival.
Conclusions: We found no significant differences in the functional outcome, pain, alignment, or survival, but a tendency toward better function using MR and better survival with GB.

Introduction

Many total knee arthroplasties (TKAs) can be implanted using 2 different surgical techniques: gap balancing (GB) and measured resection (MR). Both are widely used to achieve a well-balanced knee after a TKA [1].

The MR or bony referencing technique uses anatomical landmarks to resect an amount of the bone equal to the thickness of the prosthesis that will be implanted, taking into account wear and deformity. To prepare the femur, the distal femoral resection is performed first. The femoral rotation is set based on (at least) one of 3 bony landmarks—the posterior condylar axis, the epicondylar axis, or the anteroposterior (AP) trochlear axis (Whiteside's line)—and subsequently an adequate thickness of the bone is resected posteriorly. The tibial cut is performed at an angle of 87° to 90° with the tibial shaft. If needed, a soft-tissue release can be performed to equalize the flexion and extension gaps [2,1].

The disadvantage of MR is that the resection depends on the judgment of the surgeon to identify the bony landmarks. This has been reported to have a low reproducibility [3]. Another disadvantage is that it does not take into account the changes in laxity that can occur in flexion, after a ligamentous release is performed in extension, causing a gap mismatch.

GB uses the tension in the soft tissue to obtain a balanced flexion and extension gap. In this technique, both collateral ligaments are put under equal tension to determine the amount of the bone to be resected in flexion and in extension. As such, the femoral rotation is determined by the ligamentous tension rather than by bony landmarks.

When using this technique, the proximal tibial and distal femoral cuts are performed first. Soft tissues are then balanced in extension, resulting in a rectangular extension gap (equal, medial, and lateral gaps). Then, the knee is flexed and the joint space is

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.artd.2020.07.046.

* Corresponding author. Department of Orthopedic Surgery and Traumatology UZ Brussel, Laarbeeklaan 101, 1090 Brussels, Belgium. Tel.: +3 247 477 7997.
E-mail address: eva.dewachter@uzbrussel.be

https://doi.org/10.1016/j.artd.2020.07.046

distracted using a tensor/balancer. The posterior femoral cut is performed parallel to the tibial cut, ensuring a rectangular flexion gap, matching the extension gap [4,5].

The theoretical benefits of GB include the ability to compensate for femoral bone loss and the gap changes that occur in flexion once ligament balancing has been performed in extension.

On the other hand, using the GB technique to balance the knee in flexion and extension may raise the joint line [5]. This has been shown to cause midflexion instability and patellofemoral complications [6,7]. In addition, ligament elongation due to longstanding joint deformity could result in an asymmetric and excessive medial or lateral flexion space. As such, the femoral component could be malrotated, causing patellar maltracking [6]. Moreover, the flexion gap in a normal knee may not be truly rectangular but wider on the lateral side. This would not be reproduced with the GB technique [8]. Last but not least, a dislocated patella and chronic ligament injury or insufficiency might influence the soft-tissue tension during surgery and favor component malorientation.

The aim of this study is to compare the MR and GB techniques when using the same prosthesis (balanSys posterior-stabilized [PS], Mathys LTD Bettlach, Switzerland) in a multicenter setup without a navigation system.

Material and methods

In this multicenter retrospective study, we included patients from 4 hospitals: Dietrich Bonhoeffer Klinikum Altentreptow (Germany), Oberlinklinik Potsdam (Germany), St Josefs Krankenhaus Hilden (Germany), and Universitair Ziekenhuis Brussel (Belgium). The first 3 hospitals used the MR technique, whereas the last one used the GB technique. All implantations were performed by experienced surgeons, without the use of a navigation system. In all centers, the balanSys PS prosthesis was used. The balanSys total knee prosthesis (Mathys Ltd Bettlach, Switzerland) was first introduced in 1997 as a cruciate-retaining and later as a unicondylar system. In a second stage, a fixed-bearing PS prosthesis was developed, focusing on the natural articulation of the patella with a high congruence between the femoral component and the polyethylene inlay. The developers also tried to optimize the design of the posterior femoral condyles and the post-cam mechanism to favor a physiological femoral rollback during flexion [9].

Patients were included between January 2006 and April 2008.

Patients were asked to be evaluated in the outpatient clinic with radiographs. The patients who refused were contacted by phone and asked about pain, satisfaction, complications, and revision surgery.

Demographic data at the time of surgery (age, sex, body mass index [BMI], preoperative diagnosis) and clinical data during follow-up (Knee Society Score [KSS] and visual analog scale [VAS] for pain and satisfaction) were collected prospectively and compared between groups. We also compared the survival rate of the prosthesis in both groups, considering the revision of at least one component as a failure. Finally, we compared both groups radiographically based on anteroposterior and lateral views as well as standing full-leg radiographs. The alignment of the prosthesis was evaluated based on the angle between the mechanical axis of the femur and tibia, the femoral α angle, and the tibial β and δ angles (Fig. 1).

The α angle was measured on the medial side between a line connecting both distal femoral condyles and a line drawn along the femoral shaft axis. The β angle was measured on the medial side between a line drawn along the tibial baseplate and the tibial shaft axis. The δ angle was measured on the mediolateral view, between a line along the tibial baseplate and a line drawn along the axis of the tibial shaft on the posterior side [10].

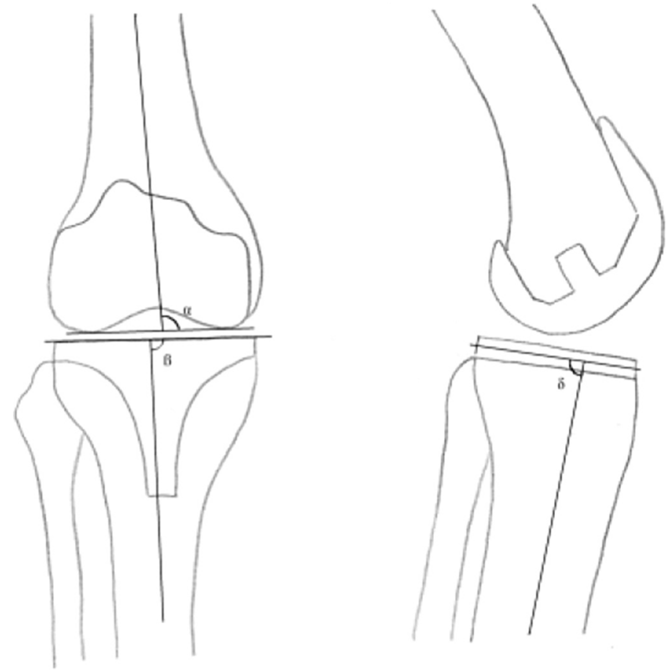


Figure 1. The alignment of the prosthesis was evaluated using the α, β, and δ angles, as described in Materials and Methods.

Radiographs taken between 1 and 3 years of follow-up were used to evaluate radiolucency and osteolysis surrounding all prosthetic components. Unfortunately, there were too few radiographs at the 7-year follow-up to perform an analysis. On the AP view, 4 femoral zones and 5 tibial zones were assessed. On the lateral view, we evaluated 4 tibial zones. First, we divided the patients into 2 groups: those with radiolucent lines ≤1 mm and those with radiolucent lines >1 mm. Then, we assigned a score to each of the 13 zones: 0 for zones without radiolucent lines and 1, 2, 3 for zones with radiolucent lines <1 mm, between 1 and 2 mm, and >3 mm, respectively. For each patient, a 'radiolucency score' was calculated by adding the scores of the 13 zones.

Interval variables (age, BMI, VAS, alignment angles) were described using the mean value and standard deviation and compared with an analysis of variance, with the different hospitals or surgical techniques as independent variables. Ordinal variables (KSS) were described using the median, the lower and upper quartiles, and the sample minimum and maximum and compared using a Kruskal-Wallis test. Nominal variables were compared using a chi-squared test. The implant survival, including patients contacted by phone, was reported by means of the Kaplan-Meier estimator.

This study was approved by the Medical Ethics Committee UZ Brussel—VUB (B.U.N. 143201419848).

Results

In total, 262 TKAs, implanted in 256 patients (6 bilateral cases), were included from 4 different centers (Appendix Table 1 and Appendix Fig. 1).

Demographics

Demographic characteristics were comparable in the GB and MR groups, except for the BMI (Appendix Table 2). Both the age and

BMI showed a significant difference between the 4 centers. Patients in the MR group were significantly younger and heavier than those in the GB group.

Function

At 2-year follow-up, there was no statistically significant difference in the knee function between the GB and MR groups as measured using the Total KSS, the Knee Score, and the Function Score. However, the Total KSS and the Function Score almost reached the level of significance between the GB and MR groups. The VAS score for pain was comparable in both groups, but the VAS score for satisfaction was significantly better in the MR group (Appendix Table 3).

At a minimum of 7-year follow-up, there was still no significant difference in the Knee Score, Function Score, Total KSS, and VAS score for pain between both groups. In contrast to 2-year follow-up, the VAS for satisfaction showed no significant difference (Appendix Table 3).

Of our 262 TKAs, 13 had a KSS of 100 or less. All 13 patients had a low preoperative knee function (median KSS: 105, median Knee score: 50, median Function score: 50). Five of these patients had other health problems that influenced their KSS. One other patient had a preoperative anatomical varus alignment of 12°. For the other 7 patients, no other preoperative reason than poor knee function could be found.

Cases with a preoperative varus (femur-tibia angle ≤3°) or valgus (femur-tibia angle ≥9°) were grouped separately to compare the Knee score, Function Score, and KSS. There was no significant difference in the Function or Knee score between these groups and patients with a neutral axis preoperatively (Appendix Table 4).

Learning curve

To evaluate the learning curve, we compared the first 15 cases of each surgeon with their following surgeries in terms of the KSS and VAS for pain and satisfaction. Overall, the learning curve was smooth and in Dietrich Bonhoeffer Klinikum Altentreptow, the KSS of the first 15 TKAs was even significantly better than that of the following surgeries. However, when looking at the Knee and Function score separately, no significant difference was found. The other KSS and VAS scores for pain and satisfaction did not differ between surgeons or between the GB and MR groups (Appendix Table 5).

Radiology

In the coronal plane, the femoral (α) and tibial angles (β) were both significantly different between the MR and GB groups. In the GB group, both angles had about 1 more degree of valgus (Appendix Table 6). That difference was not shown on full-leg standing radiographs, and all prostheses were within 2° of neutral alignment in the coronal plane. The alignment of the tibia in the sagittal plane (δ) was not significantly different.

Radiolucency

Radiolucent lines were assessed on radiographs between 1 and 3 years of follow-up. Although patients from the Oberklinik in Potsdam had significantly more radiolucent lines (Appendix Table 7) and a higher radiolucency score (Appendix Table 8), there was no significant difference between the MR and GB groups. As this same center also had a lower score for VAS satisfaction, it may point to low-grade loosening.

Revisions

Overall, of the 262 patients, 8 (3.05%) revisions were reported. Five of these revisions were performed within the first year and 3 after that (Fig. 2). All 8 revisions were performed in the MR group. Reasons for revisions are reported in Appendix Table 9.

Discussion

Comparing the MR and the GB surgical techniques to implant a balanSys PS TKA, we found no significant difference in the functional outcome, pain, and radiological alignment. Normal values for the α, β, and δ angles are not well defined in the literature. However, our results were comparable with values in other studies [11,12].

We did find a tendency toward a better functional outcome and an initial significant higher degree of satisfaction in the MR group. It is, nevertheless, questionable if that difference is enough to reach a minimal clinically important difference (MCID). In the literature, there is no consensus regarding the MCID on a VAS for satisfaction after TKA. However, after hip arthroscopy for femoroacetabular impingement, the MCID has been reported to be as high as 52.8 of 100. [13] In our study, the MR technique had a mean advantage on the VAS for satisfaction of less than 1 of 100. On the other hand, revisions were only reported in the MR and not in the GB group.

In the literature, there is little consensus regarding the risks and benefits of the GB vs the MR techniques. Two meta-analyses comparing MR and GB were found. The first included 8 studies published between 1985 and 2015. Of these 8 studies, 4 used the same prosthesis for both techniques. The second meta-analysis included 8 randomized controlled trials (RCTs) published between 1986 and 2015, of which 5 were already included in the previous meta-analysis. Of the 3 remaining RCTs, 2 used the same prosthesis for the MR and GB techniques. After 2015, 4 more studies comparing GB and MR were found, all using the same prosthesis in both groups (Appendix Table 10).

Moon et al. [14] concluded that both techniques showed similar soft-tissue balancing, with a small difference in the joint line elevation. They suggested that MR and GB techniques are not mutually exclusive. Huang et al. [2] concluded that the GB technique results in statistically significant improvement in the restoration of mechanical and rotational alignments and KSS and Function scores but resulted in a higher joint line.

Most of the previous studies used navigation systems, some only in one of the groups, while other studies used different navigation systems in both groups. The use of navigation systems could have influenced the results, and to date, many hospitals do not have access to such navigation systems. Our study did not use a navigation system,

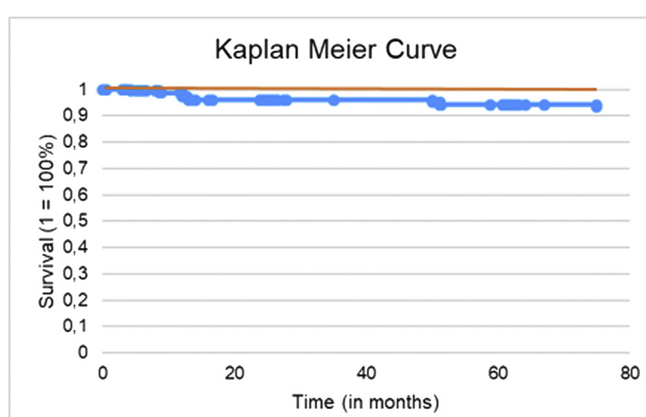


Figure 2. Survival using the Kaplan-Meier graph. Blue: MR and red: GB.

and surgeons stuck to their preferred operating technique in a multicenter setting, which represents the most common clinical situation.

Overall, we found 3 studies that implanted the same prosthesis with both surgical techniques and used no navigation system as in our setting. Compared with our study, all not only included fewer patients but also found no significant difference in the functional outcome, radiographic assessment, or revision rate for aseptic loosening. The limitations of our study include its retrospective character. As such, after 0-2 years of follow-up, 12.6% of patients were lost to follow-up and after 2-7 years, another 24.4% were lost to follow-up. Although this last number is fairly large, the clinical difference between both groups should already have been visible at the 2-year follow-up and this was not the case. Another drawback was that only one surgeon in a single center used the GB technique. This makes it difficult to assess if differences were due to the technique, the surgeon, and/or the hospital where the surgery took place. Furthermore, there was only one surgeon per center, making this a possible confounding variable. There was also no streamlining of the surgical techniques between the different centers. Concerning survival of the prosthesis, we should remark that although there seems to be a tendency toward better survival in the GB group, there are only 8 revisions in total, of which 6 are from the same hospital/surgeon. Thus, there is no clear causation. On the other hand, our average follow-up was 68.68 months, which is satisfactory to report short- to mid-term results. Concerning the radiolucent lines, one can also question the value of the data at the 2-year-follow-up, which is a moment in time without any later comparative data.

Conclusions

We found no significant difference in the functional outcome, alignment, or survival when using the MR or GB technique to implant the balanSys PS prosthesis. However, there is a tendency toward better function and clinical results in the MR group and better survival in the GB group.

Conflict of interest

T. Scheerlinck is a paid consultant for Mathys Ltd (other products), receives research support from Mathys Ltd (the product described in the article), and is a board or committee member for the European Hip Society, Belgian Hip Society, and Belgian Orthopaedic Society (BVOT); all other authors declare no potential conflicts of interest.

Acknowledgments

This research was supported by Mathys Ltd Bettlach (Switzerland). No financial support was provided. The authors would like to thank Prof. Dr. Em. Pierre-Paul Casteleyn for his contribution in the development of the balanSys PS knee arthroplasty, for initiating this study, and for performing the initial follow-up.

References

[1] Douglas DA, Komistek RD, Kim RH, Sharma A. Gap balancing versus measured resection technique for total knee arthroplasty. Clin Orthop Relat Res 2010;468(1):102.
[2] Huang T, Long Y, George D, Wang W. Meta-analysis of gap balancing versus measured resection techniques in total knee arthroplasty. Bone Joint J 2017;99-B(2):151.
[3] Babazadeh S, Dowsey MM, Stoney JD, Choong PF. Gap balancing sacrifices joint-line maintenance to improve gap symmetry: a randomized controlled trial comparing gap balancing and measured resection. J Arthroplasty 2014;29(5):950.
[4] Berger RA, Rubash HE, Seel MJ, Thomson WH, Crossett LS. Determining the rotational alignment of the femoral component in total knee arthroplasty using the epicondylar axis. Clin Orthop Relat Res 1993;(286):40.
[5] Sheth NP, Husain A, Nelson CL. Surgical techniques for total knee arthroplasty: measured resection, gap balancing, and hybrid. J Am Acad Orthop Surg 2017;25(7):499.
[6] Martin JW, Whiteside LA. The influence of joint line position on knee stability after condylar knee arthroplasty. Clin Orthop Relat Res 1990;(259):146.
[7] Konig C, Sharenkov A, Matziolis G, et al. Joint line elevation in revision TKA leads to increased patellofemoral contact forces. J Orthop Res 2010;28(1):1.
[8] Tokuhara Y, Kadoya Y, Nakagawa S, Kobayashi A, Takaoka K. The flexion gap in normal knees: an MRI study. J Bone Joint Surg Br 2004;86:1133.
[9] BalanSys® - 4in1 soft tissue-orientated surgical technique. Switserland: Mathys Medical. http://www.mathysmedical.com/index.php?id=798&L=1 [accessed 13.05.14].
[10] Hassaballa M, Budnar V, Gbejuade H, Learmonth I. Does improved instrumentation result in better component alignment in total knee arthroplasty? Orthop Rev 2011;3(1):e3.
[11] Rhee SJ, Hong SM, Suh JT. High-flexion total knee arthroplasty using NexGen LPS-Flex System: minimum 5-year follow-up results. Knee Surg Relat Res 2015;27(3):156.
[12] Chun KC, Lee SH, Baik JS, Kook SH, Han JK, Chun CH. Clinical and radiological results of cruciate-retaining total knee arthroplasty with the NexGen®-CR system: comparison of patellar resurfacing versus retention with more than 14 years of follow-up. J Orthop Surg Res 2017;12(1):144.
[13] Beck EC, Nwachukwu BU, Mehta N, et al. Defining meaningful functional improvement on the visual analog Scale for satisfaction at 2 Years after hip arthroscopy for femoroacetabular impingement syndrome. Arthroscopy 2019;36(19):734. 2.
[14] Moon YW, Kim HJ, Ahn HS, Park CD, Lee DH. Comparison of soft tissue balancing, femoral component rotation, and joint line change between the gap balancing and measured resection techniques in primary total knee arthroplasty: a meta-analysis. Medicine (Baltimore) 2016;95(39).
[15] Lee DH, Park JH, Song DI, Padhy D, Jeong WK, Han SB. Accuracy of soft tissue balancing in TKA: comparison between navigation-assisted gap balancing and conventional measured resection. Knee Surg Sports Traumatol Arthrosc 2010;18:381.
[16] Lee HJ, Lee JS, Jung HJ, Song KS, Yang JJ, Park CW. Comparison of joint line position changes after primary bilateral total knee arthroplasty performed using the navigation-assisted measured gap resection or gap balancing techniques. Knee Surg Sports Traumatol Arthrosc 2011;19:2027.
[17] Luyckx T, Peeters T, Vandenneucker H, Victor J, Bellemans J. Is adapted measured resection superior to gap-balancing in determining femoral component rotation in total knee replacement? J Bone Joint Surg Br 2012;94(9):1271.
[18] Matsumoto T, Muratsu H, Kawakami Y, et al. Soft-tissue balancing in total knee arthroplasty: cruciate-retaining versus posterior-stabilised, and measured-resection versus gap technique. Int Orthop 2014;38(3):531.
[19] Nikolaides AP, Kenanidis EI, Papavasiliou KA, Sayegh FE, Tsitouridis I, Kapetanos GA. Measured resection versus gap balancing technique for femoral rotational alignment: a prospective study. J Orthop Surg (Hong Kong) 2014;22(2):158.
[20] Sabbioni G, Rani N, Del Piccolo N, Ben Ayad R, Carubbi C, Tigani D. Gap balancing versus measured resection technique using a mobile-bearing prosthesis in computer-assisted surgery. Musculoskelet Surg 2011;95(1):25.
[21] Tigani D, Sabbioni G, Ben Ayad R, Filanti M, Rani N, Del Piccolo N. Comparison between two computer-assisted total knee arthroplasty: gap-balancing versus measured resection technique. Knee Surg Sports Traumatol Arthrosc 2010;18(10):1304.
[22] Pang HN, Yeo SJ, Chong HC, Chin PL, Ong J, Lo NN. Computer-assisted gap balancing technique improves outcome in total knee arthroplasty, compared with conventional measured resection technique. Knee Surg Sports Traumatol Arthrosc 2011;19(9):1496.
[23] Singh VK, Varkey R, Trehan R, Kamat Y, Raghavan R, Adhikari A. Functional outcome after computer-assisted total knee arthroplasty using measured resection versus gap balancing techniques: a randomised controlled study. J Orthop Surg (Hong Kong) 2012;20(3):344.
[24] Stephens BF, Hakki S, Saleh KJ, Mihalko WM. Clinical alignment variations in total knee arthroplasty with different navigation methods. Knee 2014;21(5):971.
[25] Clement ND, Makaram N, Bell J, Tiemessen CH, Mehdi SA, Livingston SJ. Colombus® computer navigated total knee arthroplasty: gap balancing versus measured resection. Knee 2017;24(6):1442.
[26] Hommel H, Kunze D, Hommel P, Fennema P. Small improvements in postoperative outcome with gap balancing technique compared with measured resection in total knee arthroplasty. Open Orthop J 2017;11:1236.
[27] Teeter MG, Perry KI, Yuan X, Howard JL, Lanting BA. Contact kinematic differences between gap balancing vs measured resection techniques for single radius posterior-stabilized total knee arthroplasty. J Arthroplasty 2017;32(6):1834.
[28] Churchill JL, Khlopas A, Sultan AA, Harwin SF, Mont MA. Gap-balancing versus measured resection technique in total knee arthroplasty: a comparison study. J Knee Surg 2018;31(1):13.

Appendices

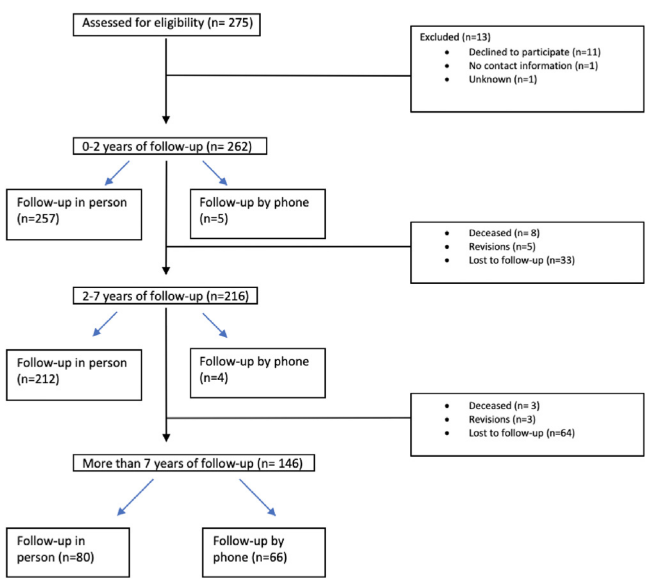
A. Appendix A1: Patients included (overview).

Appendix Table 1
Table of patients includes the ratios and follow-up time in both groups.

Technique	GB	MR			ANOVA
Center	UZ Brussel	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden	
Number of TKAs included (% of total)	100 (38.17%)	55/162	57/162	50/162	
	100 (38.17%)	162 (61.83%)			
FU time in months (median min; max 25th; 75th percentile)	70.74 ± 33.19	77.66 ± 37.20	62.03 ± 46.18	62.26 ± 30.10	$P = .07$
	70.74 ± 33.19	67.41 ± 39.19			$P = .48$

ANOVA, analysis of variance; FU, follow-up.

Appendix A2: Flowchart



Appendix Figure. Patient deceased lost to follow-up, and revised. The radiography was performed at the 2-year follow-up (at a minimum of 1 year and maximum of 3-year follow-up).

B. Appendix B1: Demographic data.

Appendix Table 2
Table of demographic characteristics.

Demographic variable	GB	MR			Chi-square/ ANOVA
	UZ Brussel	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden	
Gender M/F	30/65 (31.6%/68.4%)	19/36 (34.5%/65.5%)	24/32 (42.9%/57.1%)	18/32 (36%/64%)	Chi-square $P = .58$
	30/65 (31.6%/68.4%)	61/100 (37.9%/62.1%)			Chi-square $P = .31$
Age at operation (y) mean ± SD	71.46 ± 8.82	66.69 ± 9.54	71.08 ± 8.34	68.52 ± 9.51	ANOVA $P = .008$*
	71.46 ± 8.82	68.80 ± 9.25			ANOVA $P = .02$
BMI mean ± SD	29.38 ± 5.44	32.56 ± 6.51	29.54 ± 4.16	30.54 ± 4.63	ANOVA $P = .003$*
	29.38 ± 5.44	30.89 ± 5.34			ANOVA $P = .03$*
Preoperative diagnosis	Osteoarthrosis 98% (98), rheumatoid arthritis 0% (0), other 2% (2)	Osteoarthrosis 92.73% (51), rheumatoid arthritis 1.82% (1), other 5.45% (3)	Osteoarthrosis 98.24% (56), rheumatoid arthritis 0.18% (1), other 0% (0)	Osteoarthrosis: 98% (49), rheumatoid arthritis 2% (1), other 0% (0)	Chi-square $P = .26$
	Osteoarthrosis 97, 96% (96), rheumatoid arthritis 0.0% (0), other 2.04% (2)	Osteoarthrosis 96.30% (156), rheumatoid arthritis 1.85% (3), other 1.85% (3)			Chi-square $P = .39$

M, male; F, female; ANOVA, analysis of variance; SD, standard deviation.
The chi-squared test is used for gender and preoperative diagnosis, the ANOVA is used for the age and BMI. For gender, n = 256 (as there are 6 bilateral cases) and for the age at surgery, BMI, and diagnosis, n = 262.
* Significant difference.

Appendix B2: KSS and VAS for pain and satisfaction.

Appendix Table 3
Functional outcome at 2-year and (at least) 7-year follow-up.

Two-year follow-up		GB	MR			Statistical test
		UZ Brussel	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden	
Total KSS (0-200)	Median	169	177	187	175.5	KW
	Min; max	82; 200	20; 200	20; 200	20; 200	$P = .20$
	25th; 75th percentile	156.5; 180	161.75; 195	164.25; 196.5	155; 193.5	
		169	177			KW
		82; 200	20; 200			$P = .06$
		156.5; 180	159.75; 195			
Knee score (0-100)	Median	94	95	95	92	KW
	Min; max	48; 100	20; 100	20; 100	20; 100	$P = .24$
	25th; 75th percentile	88.75; 99	91; 97.25	92; 100	87; 97	
		94	94			KW
		48; 100	20; 100			$P = .95$
		88.75; 99	89.75; 97			
Function score (0-100)	Median	80	85	90	80	KW
	Min; max	20; 100	0; 100	0; 100	0; 100	$P = .32$
	25th; 75th percentile	70; 90	70; 100	65; 100	70; 100	
		80	80			KW
		20; 100	0; 100			$P = .07$
		70 ; 90	70; 100			
VAS pain (0-10)	Mean ± SD	1.41 ± 1.94	0.82 ± 1.91	0.76 ± 1.09	1.21 ± 0.80	ANOVA $P = .23$
		1.41 ± 1.94	0.98 ± 1.41			ANOVA $P = .10$
VAS satisfaction	Mean ± SD	8.57 ± 1.04	9.40 ± 0.82	9.47 ± 0.87	8.88 ± 0.64	ANOVA $P < .001$
		8.57 ± 1.04	9.19 ± 0.80			ANOVA $P < .001$
7-year follow-up		**GB**	**MR**			**Statistical test**
		UZ Brussel	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden	
Total KSS (0-200)	Median	160	179.5	153	176	KW
	Min; max	85; 199	144; 200	75; 199	135; 200	$P = .08$
	25th; 75th percentile	145; 197	160.75; 191	144; 166	163; 196	
		160	175			KW
		85; 199	75; 200			$P = .36$
		145; 197	157; 194.25			
Knee score (0-100)	Median	95	97	85	95	KW
	Min; max	49; 100	85; 100	55; 99	85; 100	$P = .006$
	25th; 75th percentile	90; 97	92.75; 99.25	78; 92	92; 97.5	
		95	95			KW
		49; 100	55; 100			$P = .59$
		90; 97	87.75; 99			
Function score (0-100)	Median	80	80	75	80	KW
	Min; max	35; 100	50; 100	20; 100	50; 100	$P = .41$
	25th; 75th percentile	50; 100	68.75; 100	60; 80	67.5; 100	
		80	80			KW
		35; 100	20; 100			$P = .59$
		50; 100	65; 98.5			
VAS pain (0-10)	Mean ± SD	1.18 ± 1.71	1.08 ± 2.87	2.54 ± 2.43	0.52 ± 0.79	ANOVA $P = .06$
		1.18 ± 1.71	0.98 ± 1.41			ANOVA $P = .998$
VAS satisfaction	Mean ± SD	8.59 ± 1.14	9.56 ± 0.65	8.00 ± 3.00	9.41 ± 0.83	ANOVA $P = .005$*
		8.59 ± 1.14	9.17 ± 1.67			ANOVA $P = .12$

ANOVA, analysis of variance; SD, standard deviation.
* Significant difference.

C. Appendix C: Knee Function.

Appendix Table 4
Table showing preoperative valgus and varus knees compared with preoperative neutrally aligned knees.

Score	GB				MR			
	Preoperative varus (≤3° valgus) (n = 30)	Preoperative neutral (3°-9° valgus) (n = 29)	Preoperative valgus (≥9° valgus) (n = 5)	Kruskal-Wallis test	Preoperative varus (≤3° valgus) (n = 37)	Preoperative neutral (3°-9° valgus) (n = 67)	Preoperative valgus (≥9° valgus) (n = 6)	Kruskal-Wallis test
Total KSS (0-200)	168.5	169	175	*P* = .51	184	175	191	*P* = .38
Median	146; 200	88; 200	82; 180		20; 200	20; 200	70; 200	
Min; max	162.5; 178.25	160.5; 178.25	174; 180		157; 196	159.5; 193	157; 196	
25th; 75th percentile								
Knee score (0-100)	92	94.5	95	*P* = .21	94	94	98	*P* = .22
Median	70; 100	48; 100	62; 100		20; 200	20; 100	70; 100	
Min; max	89.5; 93.5	86; 97	94; 100		87; 97	90.5; 97	95.5; 99.75	
25th; 75th percentile								
Function score (0-100)	77.5	80	80	*P* = .92	90	80	95	*P* = .46
Maedian	60; 100	30; 100	20; 80		0; 100	0; 100	0; 100	
Min; max	89.25; 93.5	70; 90	80; 80		80; 100	70; 100	71.25; 100	
25th; 75th percentile								

D. Appendix D: Learning curve.

Appendix Table 5
Table portraying the learning curve: results of the first 15 surgeries performed by each surgeon, compared with their following surgeries.

Score	GB			MR		
	First 15 TKAs	Other TKAs (n = 50)	Kruskal-Wallis test	First 15 TKAs	Other TKAs (n = 63)	Kruskal-Wallis test
Total KSS (0-200)	173	169	*P* = .65	177	177	*P* = .11
Median	146; 197	82; 200		20; 200	20; 200	
Min; max	165.5; 175	155; 180		164; 197	153; 192	
25th; 75th percentile						
Knee score (0-100)	94	94	*P* = .90	95	93	*P* = .13
Median	70; 100	48; 100		20; 100	20; 100	
Min; max	92; 95	85.75; 99		92; 98	86; 97	
25th; 75th percentile						
Function score (0-100)	80	80	*P* = .75	85	80	*P* = .27
Median	60; 100	20; 100		0; 100	0; 100	
Min; max	71.25; 80	66.25; 90		70; 100	67.5; 100	
25th; 75th percentile						
VAS pain (0-10)	1.10 ± 1.26	1.51 ± 2.10	*P* = .86	0.74 ± 0.89	1.15 ± 1.70	*P* = .32
Mean ± SD						
VAS satisfaction (0-10)	8.47 ± 0.99	8.60 ± 1.07	*P* = .44	9.28 ± 0.81	9.14 ± 0.80	*P* = .29
Median						
Min; max						
25th; 75th percentile						

SD, standard deviation.

E. Appendix E: Alignment.

Appendix Table 6
Table of radiological assessment of alignment in the coronal and sagittal planes.

Alignment		GB	MR			ANOVA
		UZ Brussel (n = 100)	Dietrich Bonhoeffer Klinikum Altentreptow (n = 54)	Oberklinik Potsdam (n = 15)	St Josefs Krankenhaus Hilden (n = 44)	
Femoral angle α (°)	Average ±SD	96.59 ± 1.39	95.50 ± 1.22	95.27 ± 3.67	95.45 ± 1.72	$P < .001$*
		96.59 ± 1.39	95.45 ± 1.88			$P < .001$*
Tibial angle β (°)	Average ±SD	91.01 ± 1.61	90.76 ± 1.15	90.47 ± 1.13	90.25 ± 0.84	$P = .02$*
		91.01 ± 1.61	90.52 ± 1.05			$P = .01$*
Tibial angle δ (°)	Average ±SD	84.03 ± 2.24	83.44 ± 3.12	85.07 ± 3.01	84.68 ± 2.09	$P = .04$*
		84.03 ± 2.24	81.14 ± 2.81			$P = .75$

The angles are described in Materials and Methods.
SD, standard deviation; ANOVA, analysis of variance.
* Significant difference.

F. Appendix F: Radiolucency.

Appendix Table 7
The patients were divided in 2 groups: without zones showing radiolucent lines > 1 mm and with zones showing radiolucent lines of >1 mm.

Radiolucent Zones	GB	MR			Chi-square
	UZ Brussel	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden	
≤ 1 mm	58/76 (76.32%)	33/50 (66%)	5/14 (35.71%)	39/44 (88.64%)	$P = .001$*
	58/76 (76.32%)	77/108 (71.30%)			$P = .45$
> 1 mm	18/76 (23.68%)	17/50 (34%)	9/14 (62.28%)	5/44 (11.36%)	$P = .001$*
	18/76 (23.68%)	31/108 (28.70%)			$P = .45$

* Significant difference.

G. Appendix 7

Appendix Table 8
Radiolucency score.

Score	GB	MR			Kruskal-Wallis test
	UZ Brussel (88)	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden	
Radiolucency score	3	3	4.5	2.5	$P = .04$*
Median	0; 10	0; 9	0; 11	0; 9	
Min; max	2; 4	2; 4	3.25; 5.75	2; 4	
25th; 75th percentile	3	3			$P = .59$
	0; 10	0; 11			
	2; 4	2; 4			

No radiolucent line = 0, line <1 mm = 1, line 1-2 mm = 2, line ≥2 mm = 3. The radiolucency score is the sum of the score of all zones on one radiograph.
* Significant difference.

Appendix Table 9
Table showing reasons for revisions were aseptic loosening, infection, joint stiffness, patellofemoral arthritis, and instability.

Reason for Revision	GB	MR		
	UZ Brussel (88)	Dietrich Bonhoeffer Klinikum Altentreptow	Oberklinik Potsdam	St Josefs Krankenhaus Hilden
Aseptic loosening	0%	1.82% (1)	0%	2% (1)
		1.30%		
Infection	0%	1.82% (1)	0%	2% (1)
		1.30%		
Knee joint stiffness (scarring and adhesions)	0%	1.82% (1)	0%	0%
		0.65%		
Patellofemoral osteoarthritis	0%	1.82% (1)	0%	0%
		0.65%		
Instability	0%	3.64% (2)	0%	0%
		1.30%		

KS, Knee Society; ROM, range of motion; QoL, quality of life.

H. Appendix H: Revisions.

I. Appendix I: Overview of the literature.

Appendix Table 10
Overview of the literature.

Meta-analysis, Moon (2016) and Huang (2017)	The same prosthesis for GB and MR?	n	Computer navigation?	Follow-up (months)	Results
Babazadeh 2014[4]	No	103	Yes	24	- Gap symmetry was significantly better using GB. - Functional outcomes and QoL: no significant difference.
Lee 2010[15]	Yes	116	Yes, but only in the GB group	Min 24 (mean, 28)	- Better outcome GB: reduced not only the postoperative alignment outlier but also the medial gap difference and achieved a more rectangular flexion and extension gap compared with MR.
Lee 2011[16]	Yes	60	Yes	Min 24	- More joint line elevation in GB - No difference in ROM, knee score, functional score
Luyckx 2012[17]	Yes	96	No	Postoperative CT	-No significant difference in rotation -No functional data
Matsumoto 2014[18]	No	1255	Yes	Min 24	-No significant difference in achieving a rectangular gap -No significant difference in the clinical outcome
Nikolaides 2014[19]	No	63	No	Postoperative CT after 7 days	-No significant difference in the femoral-component rotation -No functional data
Sabbioni 2011[20]	Yes	67	Yes	Postoperative radiograph after 4-7 days	-MR was better in preserving the joint line -No difference in coronal alignment -No difference in sagittal alignment
Tigani 2010[21]	No	126	Yes, 6 different systems	Postoperative radiograph at 4-7 months	-MR was better in the joint line preservation -No difference in the alignment or component positioning
Pang 2011[22]	Yes	140	Yes, but only in the GB group	24	-More flexion contractures of >5° in MR at a 2-year follow-up -Significant better alignment in GB -Better Function Score and Total Oxford Score at 6-month follow-up in GB -Better Total Oxford Score at 2-year follow-up in GB
Singh 2012[23]	Yes	52	Yes	24	-No functional difference
Stephens 2014[24]	No	200	Yes, but they used 2 different systems for the GB and MR groups	15	-No significant difference in the alignment (but more outliers using MR)
Studies published after 2015					
Clement 2017[25]	Yes	144	Yes	48-84 (mean, 64.8)	-GB showed a significant better Oxford Knee Score (functional outcome) -No significant difference in patient satisfaction
Hommel 2017[26]	Yes	200	Yes	120	-Slightly but a significantly better Knee Society Knee Score with GB -No significant difference in the Knee Society Function Score and Western Ontario and McMasters Universities Osteoarthritis Index (WOMAC) -No significant difference in 10-year survival
Teeter 2017[27]	Yes	24	No	12	-Similar kinematics in both groups -No significant difference in the clinical outcome (Short Form 12 [SF12], mental component score [MCS], and physical component score [PCS], Knee Society Score, and WOMAC)
Churchill 2018[28]	Yes	221	No	24-48 (mean, 36)	-No difference in revisions for aseptic loosening -No significant difference in the functional outcome (ROM, KS function, and pain score) -No difference in radiographic assessment
De Wachter 2019	Yes	252	No	22.9-34.7 (mean, 26.0)	-No significant difference in the functional outcome (Total KSS, Knee Score, Function Score, VAS for pain or satisfaction) but tendency in favor of MR -No significant difference in the alignment (in the coronal or sagittal plane) -Tendency toward a higher survival rate in the GB group

CT, computed tomography.