		一部あり　自家骨　膝蓋骨	0
		一部あり　同種骨　大腿骨	1
		一部あり　同種骨　脛骨	1
		一部あり　同種骨　膝蓋骨	0
		あり　自家骨　大腿骨	97
		あり　自家骨　脛骨	104
		あり　自家骨　膝蓋骨	0
		あり　同種骨　大腿骨	2
		あり　同種骨　脛骨	2
		あり　同種骨　膝蓋骨	0
		一部未記入	2
	生体活性材料	なし	987
	（複数回答）	一部あり　大腿骨	6
		一部あり　脛骨	4
		一部あり　膝蓋骨	0
		あり　大腿骨	8
		あり　脛骨	6
		あり　膝蓋骨	0
		未記入	14
	補強部品	なし	925
		あり	85

Ⅱ. 研究成果の刊行物・別刷

Clin Orthop Relat Res
DOI 10.1007/s11999-009-0709-x

ORIGINAL ARTICLE

Manufacturers Affect Clinical Results of THA with Zirconia Heads

A Systematic Review

Hiroyuki Yoshitomi MD, PhD, Satoru Shikata MD, PhD, Hiromu Ito MD, PhD, Takeo Nakayama MD, PhD, Takashi Nakamura MD, PhD

Received: 30 June 2008 / Accepted: 9 January 2009

Abstract In the 1980s, zirconia was introduced for THA with the expectation of lower polyethylene wear and better clinical results. However, several studies have reported poor survivorship of zirconia-polyethylene THA. We performed a systematic review and meta-analysis of zirconia-polyethylene THA to confirm or refute the theoretical advantages of this combination. Of 163 studies identified by a comprehensive search, seven met our selection criteria. These involved 769 hips of 586 patients with a mean age of 56.8 years and a minimum followup of 60 months (mean, 89.2 months; range, 60–155 months). The consolidated revision rate of zirconia-polyethylene THA at 89.2 months was higher than that of nonzirconia-polyethylene THA by 5% (risk difference, 0.05; 95% confidence interval, 0.02–0.08). Subgroup meta-analysis suggested THAs with zirconia heads from Ceraver had more revision surgery than nonzirconia heads (risk difference, 0.08; 95% confidence interval, 0.03–0.14), whereas zirconia heads from DePuy did not (risk difference, 0.02; 95% confidence interval, −0.01–0.06). The meta-analysis for annual linear polyethylene wear (which did not involve zirconia heads from Ceraver because of insufficient descriptions) showed no difference between zirconia and control groups. Collectively, THAs with high-quality zirconia heads appear to have prosthesis survivorship and polyethylene wear equivalent to those of THAs with traditional materials, but differing quality among zirconia heads could lead to poor survivorship of prostheses.
Level of Evidence: Level III, therapeutic study. See the Guidelines for Authors for a complete description of levels of evidence.

One or more of the authors (HY, HI, TN, TN) have received funding from Health and Labour Sciences Research Grants.

H. Yoshitomi (✉), H. Ito, T. Nakamura
Department of Orthopaedic Surgery, Kyoto University Graduate School of Medicine, 54 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan
e-mail: yositomi@kuhp.kyoto-u.ac.jp

S. Shikata
Department of Surgery, Soseikai General Hospital, Kyoto, Japan

T. Nakayama
Department of Health Informatics, Kyoto University Graduate School of Medicine, Kyoto, Japan

Published online: 31 January 2009

Introduction

Ultrahigh-molecular-weight polyethylene (UHMWPE) acetabular components provide one of the most promising bearing surfaces in THA [18], and almost one million UHMWPE components are used worldwide annually [18]. Although better implant design and surgical techniques have improved the clinical results of THA, polyethylene wear debris remains a major cause of periprosthetic osteolysis and aseptic loosening, leading to revision surgery. UHMWPE acetabulum and metal heads such as stainless steel or cobalt-chromium (Co-Cr) are well-established combinations for the bearing surface used in THA.

Alumina ceramic heads were introduced in the early 1970s and had low polyethylene wear and revision rates [20]. However, they are brittle and can fracture, with an incidence ranging from 0.02% to 0.14% [3, 9, 13, 25]. Tougher and stronger zirconia ceramics were introduced clinically in the early 1980s. These produced lower polyethylene wear in vitro, equaling that of alumina ceramics [4, 17, 26]. Despite such promising data, one report suggests poor survivorship of zirconia-polyethylene THA, with 14% of aseptic loosening at a mean followup of

5.8 years [1], and another reported intensive phase transformation in the crystal structure of retrieved zirconia head ranging from 20% to 30% [10], which can lead to increased roughness of the zirconia head [26], increased polyethylene wear [3], and poor survivorship of the prosthesis [3].

We conducted a meta-analysis to (1) compare the survivorship of zirconia-polyethylene THA with that of nonzirconia-polyethylene THA, (2) ascertain interstudy heterogeneity, (3) determine whether manufacturers or fixation method influenced survivorship, and (4) determine whether there was any difference in mean annual polyethylene wear between zirconia and nonzirconia groups.

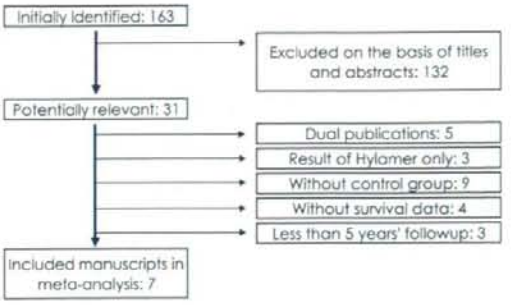

Fig. 1 A flowchart illustrates how we selected the seven studies included in our meta-analysis according to our inclusion and exclusion criteria.

Materials and Methods

We searched for reports of clinical trials that compared THA using zirconia heads combined with UHMWPE with those using other head materials. The studies needed to have followups more than 5 years regardless of the size of the femoral head used or whether cemented or cementless arthroplasty was used. The study designs comprised randomized controlled trials (RCTs) and nonrandomized trials (non-RCTs), including cohort studies and historical cohort studies. We excluded uncontrolled case series (ie, Level IV) in the meta-analysis, because it is difficult to make definitive and reproducible criteria for nonzirconia studies that criteria covers the duration of followup, background of patients, fixation method and type of implants, and year of surgery. We initially searched PubMed (1966 to July 2007), EMBASE (1974 to July 2007), and the Cochrane Central Register of Controlled Trials (Issue 4, July 2007) with the Boolean operators "zirconia AND (head OR arthroplasty OR replacement)" without MEDLINE field tags. The Japan Centra Revuo Medicina also was searched for such articles written in Japanese (1983 to July 2007). Of 163 articles identified by our initial search, one article was in French and three in German, both with English abstracts; 132 were in English; and 27 were in Japanese with Japanese abstracts. We excluded irrelevant studies if the title and abstract indicated obviously irrelevant topics, such as dental materials, knee arthroplasty, and basic research including animal experiments, biomaterials, and simulators. We excluded dual publication, uncontrolled studies, studies without the number of revised cases, clinical studies that covered only enhanced UHMWPE (Hylamer^{TM}), which produced disastrous clinical results compared with UHMWPE [24], or if they only discussed ceramic-on-ceramic THA (Fig. 1).

One reviewer (HY) extracted the age, gender, and diagnosis of patients; duration of followup; fixation method and type of implants; number of revised and unrevised cases; and the mean, standard deviation, and measurement method of annual linear polyethylene wear in groups receiving THA with either zirconia or nonzirconia heads. Another reviewer (HI) checked the accuracy of data and disagreements were resolved by discussion. Cases that were revised because of infection or fracture were removed. Revision surgeries attributable to polyethylene failure were not described in the included studies.

Seven studies (four non-RCTs and three RCTs) published in English met our inclusion criteria providing clinical results, including survivorship of zirconia-polyethylene and nonzirconia-polyethylene THA with more than 5 years' followup. These seven studies involved 769 hips of 586 patients with a mean age of 56.8 years and a minimum followup of 60 months (mean, 89.2 months; range, 60–155 months). Of the THAs, 57.0% were in women, 53.3% were for osteoarthritis, 33.3% for aseptic necrosis of the femoral head, and 6.4% for rheumatoid arthritis [1, 11, 14–16, 21, 30]. Control materials of femoral heads consisted of stainless steel, Co-Cr, and alumina ceramics.

Two reviewers (HY, HI) independently used a checklist of 11 items or 40 items to evaluate the internal validity or the generalizability of included studies [22, 29]. Disagreements were resolved by discussion. The mean scores for the four non-RCTs and three RCTs were 4.5 (range, 4–5) and 6.3 (range, 6–7), respectively, for internal validity and 19 (range, 18–20) and 21.3 (range, 21–22), respectively, for generalizability.

Two of the four non-RCT studies used historical cohorts [1, 21], and one of these cited historical results of alumina head THA in the same institute published in another study [20] (Table 1). The other two non-RCT studies were cohort studies [11, 14]. Two studies were RCTs [15, 30]. In one trial, 50 patients received bilateral THAs with a zirconia head in one hip and a Co-Cr head in the other; 48 patients had a unilateral THA with a Co-Cr head [16]. Prosthesis survival data from the 50 patients with bilateral THAs were

Table 1. Summary of analyzed studies

Study	Control materials	Head diameter (mm)	Fixation	Design	Number of zirconia	Number of control	Manufacturer of zirconia
Allain et al. [1] (1999)	Alumina	28, 32	Cement	HC	78	117	Ceraver
Hernigou and Bahrami [11] (2003)	Alumina, stainless steel	28, 32	Cement	Cohort	40	96	Ceraver
Inoue et al. [14] (2006)	Co-Cr	22	Cement	Cohort	13	13	Nippon Tokushu Tougyou
Liang et al. [21] (2007)	Alumina	22	Cement	HC	58	46	JMM
Kim [15] (2005)	Co-Cr	28	Cement, cementless	RCT	52	52	DePuy
von Schewelov et al. [30] (2005)	Stainless steel	22	Cement	RCT	28	28	DePuy
Kim et al. [16] (2003)	Co-Cr	22	Cementless	*	50	98	DePuy

* Survival data from 50 patients with bilateral THAs were treated as a RCT and the linear wear rate from 98 patients including bilateral and unilateral THAs was treated as a non-RCT; HC = studies with historical cohort; RCT = randomized controlled trial.

Table 2. Details of polyethylene

Study	Material	γ irradiation	Manufacturer	Measurement of wear
Allain et al. [1] (1999)	UHMWPE	No information	Ceraver	Film
Hernigou and Bahrami [11] (2003)	UHMWPE GUR415	γ in air	Ceraver	Digitized radiograph
Inoue et al. [14] (2006)	UHMWPE with MW > 5500 kDa*	No γ irradiation	Mizuho	Computer aid
Liang et al. [21] (2007)	UHMWPE GUR402	2.5 Mrad γ in air	JMM	Computer aid
Kim [15] (2005)	UHMWPE ram-extruded GUR1050	2.4–4 Mrad γ in vacuum	DePuy	Computer aid
von Schewelov et al. [30] (2005)	UHMWPE GUR415	γ in air	DePuy	Film
Kim et al. [16] (2003)	UHMWPE ram-extruded GUR1050*	2.4–4 Mrad γ in air*	DePuy	Computer aid

* Information obtained directly from authors; UHMWPE = ultrahigh-molecular-weight polyethylene; MW = molecular weight.

treated as a RCT, and the linear wear rate of polyethylene from all 98 patients was treated as a non-RCT. All seven studies provided the number of cases needing revision after zirconia and nonzirconia THA. Acetabular and femoral components were fixed by cement in five studies [1, 11, 14, 21, 30], whereas they were fixed by a cementless method in one study [16]. In one study, all the acetabular components were cementless, but the femoral components were cemented or cementless [15]. The zirconia heads came from four manufacturers: Ceraver (Roissy, France), Nippon Tokushu Tougyou (Nagoya, Japan), JMM (Kobe, Japan), and DePuy (Leeds, United Kingdom). Three studies were conducted in Europe and four in Asia. The studies from Europe consisted of two studies for Ceraver and one for DePuy, and the studies from Asia consisted of one for JMM, one for Toyo Tokushu Tougyou, and two for DePuy. All polyethylenes used in the included studies were traditional UHMWPE (Table 2). Although the details of UHMWPE in one study were unavailable [1], the material and sterilization of UHMWPE in the zirconia and nonzirconia groups in the remaining studies were the same. UHMWPE was provided by Ceraver, Mizuho (Tokyo, Japan), JMM, or DePuy. The annual linear polyethylene wear was measured directly on films in two studies, on digitized radiographs in one study, and with computer aid in four studies (Table 2). Two studies using zirconia from Ceraver did not provide the standard deviation of annual linear polyethylene wear [1, 11]. The annual linear polyethylene wear rates of the remaining five studies were combined [14–16, 21, 30].

For the analysis of the risk of revision surgery, we used the fixed-effects model weighted by the Mantel-Haenszel method following a test of heterogeneity [23]. The quantity I^2 was used for assessing heterogeneity between trials in meta-analysis. A value of I^2 greater than 50% indicates substantial heterogeneity [12]. If the hypothesis of heterogeneity was accepted, we used a random-effects model, the DerSimonian-Laird method [5]. In five of seven studies for this meta-analysis, there were no revised cases in the control group [14–16, 21, 30]. No events in an outcome were considered a zero cell in the 2×2 table. Zero cells create problems in calculation of odds or risk ratio caused by the division by zero. Although this problem is commonly dealt with by adding 0.5 to each cell of the 2×2 table, this addition causes inaccuracy of odds and risk ratios if there is a high frequency of zero cells [7].

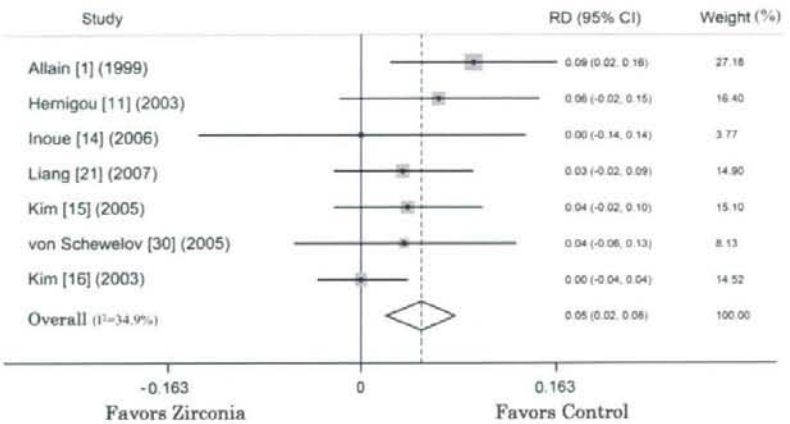

Fig. 2 When seven studies were pooled in a fixed-effects model, the risk difference (RD) of revision surgery between zirconia-polyethylene and nonzirconia-polyethylene THAs was 0.05 (95% CI, 0.02–0.08).

Therefore, we estimated the risk difference of revision surgery with 95% confidence intervals (CIs) between zirconia and control groups instead of odds or risk ratio. The risk difference of revision surgery indicates the difference of the frequency of revision surgery during followup between groups. Furthermore, we analyzed subgroups consisting of the manufacturers of zirconia or the fixation methods of implants. To determine the existence of heterogeneous studies, an informal graphic exploration was performed using a L'Abbe plot, and sensitivity analysis was performed [19, 28]. The combined mean difference of annual linear polyethylene wear between zirconia-polyethylene and nonzirconia-polyethylene THAs was calculated according to the empirical Bayes method [27]. A random-effects model was used if I^2 was greater than 50%. If not, a fixed-effects model was used.

Studies with significant differences or those with generally expected results tend to be submitted and accepted, leading to publication bias in meta-analyses. The funnel plot, Begg's test, and Egger's test were used to evaluate the potential for publication bias (ie, for better or worse survivorship) associated with the survivorship of THA [2, 8]. A p value of publication bias less than 0.10 was considered significant. The funnel plot was not symmetric and both tests revealed publication bias (Begg's test, $p = 0.004$; Egger's test, $p = 0.000$), indicating there might be unpublished studies with insignificant results or unexpected results.

We used the Stata® software package (Version 9.2; StataCorp LP, College Station, TX) for all analyses.

Results

Zirconia-polyethylene THAs required more revision surgery at a mean followup of 89.2 months (range, 60–155 months) than nonzirconia-polyethylene THAs (Fig. 2). The estimated pooled risk difference of revision surgery from seven studies in the fixed-effects model was 0.05 (95% CI, 0.02–0.08; $p = 0.001$; homogeneity $I^2 = 34.9\%$). The pooled risk difference estimated from the three RCT studies suggests similar risk of revision at a mean followup of 113.2 months with zirconia heads and nonzirconia heads (risk difference, 0.02; 95% CI, −0.01–0.06; $p = 0.221$; homogeneity $I^2 = 0.0\%$).

We identified the study by Allain et al. [1] as the source of heterogeneity. That study used so-called first-generation zirconia ceramics, which are believed to provide inferior clinical results [3]. The combined revision rate of zirconia-polyethylene THA at a mean followup of 108.1 months estimated from the remaining six studies also was greater than that of nonzirconia-polyethylene THA by 3.4% (risk difference, 0.034; 95% CI, 0.003–0.06; $p = 0.03$) and interstudy heterogeneity of the remaining six studies was similar (homogeneity, $p = 0.03$).

THAs with zirconia heads made by Ceraver led to more revision surgery than control materials (risk difference, 0.08; 95% CI, 0.03–0.14; $p = 0.02$; homogeneity $I^2 = 0.0\%$), whereas THA with zirconia heads from DePuy did not (risk difference, 0.02; 95% CI, −0.01–0.06; $p = 0.212$; homogeneity $I^2 = 0.0\%$). These results suggest there are differences in clinical results between manufacturers of zirconia probably because of the quality of the zirconia heads.

To assess the influence of fixation methods of implants on the survivorship of THA with zirconia, we excluded each study observing only cementless THAs [16] or observing cemented and cementless THAs [15]. Cemented THAs with zirconia heads had more revision surgery than cemented THAs with control materials (risk difference, 0.06; 95% CI, 0.02–0.10; $p = 0.001$; homogeneity $I^2 = 2.8\%$). This result implies fixation methods of implants

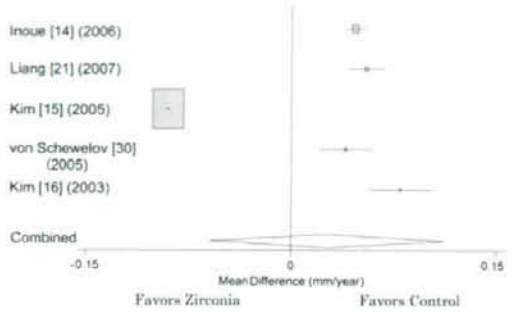

Fig. 3 When five studies providing the means and standard deviations of polyethylene wear rate were pooled in a random-effects model, the difference in annual linear polyethylene wear between zirconia-polyethylene and nonzirconia-polyethylene THAs was −0.023 mm per year (95% CI, −0.120–0.074).

have little influence on the difference of survivorship between zirconia and nonzirconia groups.

Five studies providing the means and standard deviations of annual linear polyethylene wear did not include studies using zirconia of Ceraver (Table 2). There was no difference in the pooled annual polyethylene wear between zirconia head groups and control groups from these five studies. The pooled difference of annual polyethylene wear in the random-effects model was 0.026 mm per year (95% CI, −0.060–0.113; p = 0.552; homogeneity $I^2 = 99\%$) (Fig. 3).

Discussion

Zirconia was introduced for THA in the 1980s with the expectation of lower polyethylene wear and better clinical results. However, several studies have reported poor survivorship of zirconia-polyethylene THAs [1, 10]. In this study, we conducted a systematic review and a meta-analysis to assess the survivorship and polyethylene wear of zirconia-polyethylene THA compared with nonzirconia-polyethylene THA and the influence of zirconia manufacturers or implant fixation method on the survivorship.

Some study limitations of this meta-analysis should be considered. First, cross-linked polyethylene was available from 1998 and polyethylene for THA had almost shifted from traditional UHMWPE to cross-linked UHWMWPE because of low polyethylene wear [6]. For our meta-analysis, only clinical results of THA with traditional UHMWPE were available. Although full translation of this meta-analysis to THA with cross-linked UHMWPE is difficult, it is reasonable to anticipate head materials causing high wear of traditional UHMWPE and poor survivorship of THA also can cause high wear of cross-linked UHMWPE and poor clinical results compared with those of other materials. Therefore, we need to clarify the clinical outcomes of THA with a zirconia head to determine whether future use of zirconia heads is justified. Second, bias in material or sterilization methods of traditional UHMWPE or geographic area where studies were conducted might exist. Several kinds of materials or sterilization methods of traditional UHMWPE were used in the reported studies. Two studies using zirconia of Ceraver were conducted in France, whereas of the three studies of DePuy, one was conducted in Europe and two in Asia. However, the main difference between zirconia and nonzirconia groups in each study was head materials. Therefore, we considered it justified to consolidate the difference of risk of revision or linear polyethylene wear in the zirconia and nonzirconia groups among studies. Third, our analysis indicated the possibility of publication bias in the clinical results of zirconia-polyethylene THA. Thus, there might have been studies that remained unpublished for various reasons, which could not be included in our meta-analysis. If one considers a possible publication bias of this analysis toward worse clinical results for zirconia-polyethylene THA, unpublished studies, in theory, might have indicated better clinical outcomes. However, such studies would have been published actively, because zirconia was expected to provide clinical results as good as those seen in vitro. Another possibility is that revision rates of zirconia-polyethylene THAs were disastrous and surgeons simply stopped using a device without studying the differences. However, disastrous clinical results, even if in an unexpected direction, tend to be submitted and accepted, as in the cases of early zirconia heads [1, 10] or Hylamer™ [24]. Therefore, unpublished studies that would have shifted the meta-analysis toward better clinical results for zirconia are unlikely. Meta-analyses can use only peer-reviewed published data. The International Committee of Medical Journal Editors proposed a system of clinical trial registration in 2004, and many journals support this system. In the future, more RCTs examining the clinical results of zirconia heads are expected to be registered and published regardless of their results, and further meta-analyses involving such trials are expected.

Although the most likely cause of the poor survival of THAs with zirconia heads is polyethylene wear, we found no difference between zirconia and control groups. This might be attributed to the following factors. First, only five of seven studies provided necessary data for the consolidation of annual linear polyethylene wear [14–16, 21, 30]. Second, the data for zirconia from Ceraver, which considerably increased revision surgery of THA, was not involved in the meta-analysis for polyethylene wear [1, 11]. Hernigou and Bahrami [11] reported mean annual polyethylene wear rates against zirconia from Ceraver and

stainless steel of 0.4 mm per year and 0.13 mm per year, respectively. Allain et al. [1] reported a mean annual linear polyethylene wear rate against zirconia from Ceraver in the revised cases of 0.5 mm per year, whereas total mean polyethylene wear rates against zirconia and alumina were 0.09 mm per year and 0.1 mm per year, respectively. Their results indicate heterogeneous quality among zirconia heads. Third, the polyethylene wear rate against zirconia reported by Kim [15] in 2005 was low with a narrow CI. This might cancel the other four studies, which favor the control group. Collectively, THAs with high-quality zirconia heads may have prosthesis survivorship and polyethylene wear equivalent to those of THAs with traditional materials, but heterogeneous quality among zirconia heads leads to poor survivorship of the prosthesis because of polyethylene wear.

Transformation from the tetragonal to a monoclinic phase of yttrium-stabilized tetragonal zirconia polycrystals leads to volumetric expansion of approximately 3% to 4% [26], increased roughness of the zirconia head [26], and increased polyethylene wear [3]. Although zirconia manufacturers predicted there would be less than 2% monoclinic sites on zirconia femoral heads during the first 10 years in patients, one retrieval study suggests transformation of as much as 80% by 9 years [3]. That the area of monoclinic transformation corresponds to the contact region against polyethylene indicates tribologic conditions also trigger phase transformation. Hydrothermal conditions such as autoclaving for sterilization also are involved in phase transformation. The manufacturing process, manufacturers, and year of manufacture can be crucial for stability of the tetragonal phase in zirconia. The first generation of zirconia from some manufacturers has undergone substantial transformation of its crystal structure to a monoclinic phase and these implants have inferior survivorship [1, 10]. Our subgroup meta-analysis supports the importance of the manufacturing process and manufacturers. Studies using zirconia from DePuy showed no difference in risk difference between the zirconia and control heads. This implies zirconia heads from specific manufacturers could give clinical results at least equivalent to those of traditional head materials.

Should we now stop using zirconia heads? We believe urgent cessation of zirconia-polyethylene THA would not be justified, because the zirconia supplied by some manufacturers seems to have clinical results equivalent to those of traditional materials. Obviously, alumina ceramics had clinically lower polyethylene wear than conventional metal heads and better ceramics with toughness and low polyethylene wear desired. Yttrium-stabilized tetragonal zirconia polycrystals with 0.25% (w/w) Al_2O_3 (Al-doped) has been available since 2002. Al-doped zirconia ceramics are approximately fivefold more resistant to phase transformation by hydrothermal conditions than non-Al-doped zirconia and are expected to have lower polyethylene wear and provide better clinical results in THA [3, 21]. Additional clinical confirmation of trials for zirconia ceramics, including such improved materials in combination with cross-linked UHMWPE, by meta-analyses of published and registered studies is required.

Acknowledgments We thank Keiichi Kawanabe for discussion and Koji Asaumi and Young-Hoo Kim for providing information about polyethylene.

References

1. Allain J, Le Mouel S, Goutallier D, Voisin MC. Poor eight-year survival of cemented zirconia-polyethylene total hip replacements. *J Bone Joint Surg Br.* 1999;81:835–842.
2. Begg C, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 1994;50:1088–1101.
3. Brown S, Green D, Pezzotti G, Donaldson T, Clarke I. Possible triggers for phase transformation in zirconia hip balls. *J Biomed Mater Res B Appl Biomater.* 2008;85:444–452.
4. Cales B. Zirconia as a sliding material: histologic, laboratory, and clinical data. *Clin Orthop Relat Res.* 2000;379:94–112.
5. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7:177–188.
6. Digas G, Karrholm J, Thanner J, Herberts P. 5-year experience of highly cross-linked polyethylene in cemented and uncemented sockets: two randomized studies using radiostereometric analysis. *Acta Orthop.* 2007;78:746–754.
7. Egger M, Altman DG, Smith GD. *Systematic Reviews in Health Care: Meta-Analysis in Context.* 2nd ed. London, UK: BMJ Books; 2001.
8. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315:629–634.
9. Fritsch E, Gleitz M. Ceramic femoral head fractures in total hip arthroplasty. *Clin Orthop Relat Res.* 1996;328:129–136.
10. Haraguchi K, Sugano N, Nishii T, Miki H, Oka K, Yoshikawa H. Phase transformation of a zirconia ceramic head after total hip arthroplasty. *J Bone Joint Surg Br.* 2001;83:996–1000.
11. Hernigou P, Bahrami T. Zirconia and alumina ceramics in comparison with stainless-steel heads: polyethylene wear after a minimum ten-year follow-up. *J Bone Joint Surg Br.* 2003;85:504–509.
12. Higgins J, Thompson S, Deeks J, Altman D. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557–560.
13. Higuchi F, Shiba N, Inoue A, Wakebe I. Fracture of an alumina ceramic head in total hip arthroplasty. *J Arthroplasty.* 1995;10:851–854.
14. Inoue A, Asaumi K, Endo H, Fujiwara K, Mitani S, Ozaki T. Assessment of head wear more than ten years after total hip arthroplasty: 22-mm zirconia vs. metal heads. *Acta Med Okayama.* 2006;60:311–318.
15. Kim YH. Comparison of polyethylene wear associated with cobalt-chromium and zirconia heads after total hip replacement: a prospective, randomized study. *J Bone Joint Surg Am.* 2005;87:1769–1776.
16. Kim YH, Oh SH, Kim JS, Koo KH. Contemporary total hip arthroplasty with and without cement in patients with osteonecrosis of the femoral head. *J Bone Joint Surg Am.* 2003;85:675–681.

17. Kumar P, Oka M, Ikeuchi K, Shimizu K, Yamamuro T, Okumura H, Kotoura Y. Low wear rate of UHMWPE against zirconia ceramic (Y-PSZ) in comparison to alumina ceramic and SUS 316L alloy. *J Biomed Mater Res.* 1991;25:813–828.
18. Kurtz S, Muratoglu O, Evans M, Edidin A. Advances in the processing, sterilization, and crosslinking of ultra-high molecular weight polyethylene for total joint arthroplasty. *Biomaterials.* 1999;20:1659–1688.
19. L'Abbé K, Detsky A, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med.* 1987;107:224–233.
20. Le Mouel S, Allain J, Goutallier D. [10-year actuarial analysis of a cohort of 156 total hip prostheses of a cemented polished aluminum/polyethylene alloy] [in French]. *Rev Chir Orthop Reparatrice Appar Mot.* 1998;84:338–345.
21. Liang B, Kawanabe K, Ise K, Iida H, Nakamura T. Polyethylene wear against alumina and zirconia heads in cemented total hip arthroplasty. *J Arthroplasty.* 2007;22:251–257.
22. Malmivaara A, Koes BW, Bouter LM, van Tulder MW. Applicability and clinical relevance of results in randomized controlled trials: the Cochrane review on exercise therapy for low back pain as an example. *Spine.* 2006;31:1405–1409.
23. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst.* 1959; 22:719–748.
24. Norton M, Yarlagadda R, Anderson G. Catastrophic failure of the Elite Plus total hip replacement, with a Hylamer acetabulum and zirconia ceramic femoral head. *J Bone Joint Surg Br.* 2002; 84:631–635.
25. Peiró A, Pardo J, Navarrete R, Rodriguez-Alonso L, Martos F. Fracture of the ceramic head in total hip arthroplasty: report of two cases. *J Arthroplasty.* 1991;6:371–374.
26. Piconi C, Maccauro G. Zirconia as a ceramic biomaterial. *Biomaterials.* 1999;20:1–25.
27. Robbins H. *An Empirical Bayes Approach to Statistics, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability.* Vol 1. Berkeley, CA: University of California Press; 1956.
28. Sharp S, Thompson S, Altman D. The relation between treatment benefit and underlying risk in meta-analysis. *BMJ.* 1996;313: 735–738.
29. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine.* 1997;22: 2323–2330.
30. von Schewelov T, Sanzen L, Onsten I, Carlsson A, Besjakov J. Total hip replacement with a zirconium oxide ceramic femoral head: a randomised roentgen stereophotogrammetric study. *J Bone Joint Surg Br.* 2005;87:1631–1635.

Can Magnetic Resonance Imaging–Derived Bone Models Be Used for Accurate Motion Measurement with Single-Plane Three-Dimensional Shape Registration?

Taka-aki Moro-oka,[1,2,3] Satoshi Hamai,[3] Hiromasa Miura,[3] Takeshi Shimoto,[4] Hidehiko Higaki,[4] Benjamin J. Fregly,[1] Yukihide Iwamoto,[3] Scott A. Banks[1,5]

[1]Department of Mechanical & Aerospace Engineering, University of Florida, 318 MAE-A, Mail Stop 116250, Gainesville, Florida 32611-6250

[2]Seimeikai Moro-oka Orthopaedic Hospital, Fukuoka, Japan

[3]Department of Orthopaedic Surgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

[4]Department of Mechanical Engineering, Faculty of Engineering, Kyushu Sangyo University, Fukuoka, Japan

[5]The BioMotion Foundation, Palm Beach, Florida

Received 30 March 2006; accepted 8 November 2006

Published online 8 February 2007 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/jor.20355

ABSTRACT: The purpose of this study was to compare three-dimensional (3D) kinematic measurements from single-plane radiographic projections using bone models created from magnetic resonance imaging (MRI) and computed tomography (CT). MRI is attractive because there is no ionizing radiation, but geometric field distortion and poor bone contrast degrade model fidelity compared to CT. We created knee bone models of three healthy volunteers from both MRI and CT and performed three quantitative comparisons. First, differences between MRI– and CT–derived bone model surfaces were measured. Second, shape matching motion measurements were done with bone models for X-ray image sequences of a squat activity. Third, synthetic X-ray images in known poses were created and shape matching was again performed. Differences in kinematic results were quantified in terms of root mean square (RMS) error. Mean differences between CT and MRI model surfaces for the femur and tibia were −0.08 mm and −0.14 mm, respectively. There were significant differences in three of six kinematic parameters comparing matching results from MRI–derived bone models and CT–derived bone models. RMS errors for tibiofemoral poses averaged 0.74 mm for sagittal translations, 2.0 mm for mediolateral translations, and 1.4° for all rotations with MRI models. Average RMS errors were 0.53 mm for sagittal translations, 1.6 mm for mediolateral translations, and 0.54° for all rotations with the CT models. Single-plane X-ray imaging with model-based shape matching provides kinematic measurements with sufficient accuracy to assess knee motions using either MRI– or CT–derived bone models. However, extra care should be taken when using MRI–derived bone models because model inaccuracies will affect the quality of the shape matching results. © 2007 Orthopaedic Research Society. Published by Wiley Periodicals, Inc. J Orthop Res 25:867–872, 2007

Keywords: magnetic resonance imaging (MRI); image distortion; computed tomography (CT); three-dimensional kinematics; knee

INTRODUCTION

Shape matching techniques have been used for 15 years to determine knee arthroplasty motions from fluoroscopic image sequences.[1–3] Recently, these techniques have been applied for motion measurement in joints without metallic implants, where three-dimensional (3D) surface models of the bones are created from magnetic resonance imaging (MRI)[4] and computed tomography (CT).[2,5–7] However, we are unaware of any rigorous assessment of the use of MRI–derived models for the purpose of shape registration–based motion measurement. DeFrate and colleagues reported the advantages of MRI-based model creation included the ability to add cartilage to the models and to avoid radiation exposure, but they did not directly assess the accuracy of their

Correspondence to: Scott A. Banks (Telephone: 352-392-6109; Fax: 352-392-7303; E-mail: banks@ufl.edu)

shape matching technique using the MRI–derived bone models.[4,8] It is well recognized that MRI provides lower bone contrast than CT and suffers from spatial distortions which vary by scanner, scan sequence, and the object being scanned.[9–18] MRI use is limited, for example, in stereotactic surgery of the brain unless geometric distortion correction has been performed.[14,15] Similarly, shape registration–based motion measurement requires submillimeter model accuracy for many clinically relevant measurement scenarios. In contrast, CT has negligible scaling error because images are reconstructed from line-of-sight X-ray optics.[9,11,13] The purposes of this study were to compare 3D kinematics from model-based shape matching using CT– and MRI–derived bone models and to determine if MRI–derived bone models provide sufficient fidelity to provide clinically relevant measurements.

METHODS

Three healthy subjects gave informed consent to participate in this study as approved by the institutional review board. Geometric bone models of the femur and tibia/fibula were created from CT (Toshiba, Aquilion, Tochigi, Japan) and MRI (Hitachi, Airis II Comfort, 0.3 T, Tokyo, Japan) scans of one leg. CT scans used a 512×512 image matrix, a 0.35×0.35 pixel dim, and a 1.00-mm thickness spanning approximately 150 mm above and below the joint line of the knee, and 2-mm slices through the centers of the hip and ankle joints. CT scan time was 49 s. MRI scans used a 512×512 image matrix, a 0.39×0.39 pixel dim, and a 1.0-mm thickness spanning more than 80 mm above and below the joint line of the knee. The MRI protocol was 3DT1GE, RF spoiled SARGE(RSSG). MR scan time ranged from 11 to 15 min.

Exterior cortical bone edges were segmented using commercial software (SliceOmatic, Tomovision, Montreal, CA), and these point clouds were converted into polygonal surface models (Geomagic Studio, Raindrop Geomagic, Research Triangle Park, NC). Interior cortical bone edges were not included because of poor definition in the epiphyseal and metaphyseal regions.

Anatomical coordinate systems were embedded in each bone model following a combination of previous approaches.[5,19,20] The coordinate systems were first defined for the CT models. The mediolateral (x) axes of the femur and tibia/fibula were defined by fitting a cylinder to each posterior condyle of the femur. The midpoint of the cylindrical axis was defined as the origin. The proximal/distal (y) axis for the femur was defined by a line perpendicular to the cylindrical axis in the plane intersecting the femoral head center. The proximal/distal (y) axis for the shank was perpendicular to the cylindrical axis in the plane intersecting the ankle center. The anteroposterior (z) axis was formed from the cross product of the first two. Next, the MRI model was registered with the corresponding CT model in its initial reference pose to align the embedded coordinate systems in each bone model. Automated alignment software was used with a proprietary algorithm to match 3D surfaces (Geomagic Studio). Fitting results were accurate to less than 0.1 microns in length and 0.1 arc seconds (1/36,000 of a degree) in angle compared to the official reference value.

The CT models were then shortened to the same length as the MRI models. Three experiments were performed to compare the CT and MR models.

Experiment 1

Differences between MRI and CT model surfaces were measured using inspection software (Geomagic Studio).

Experiment 2

Shape matching with the CT– and MRI–derived models was performed independently, without any information from one kinematic solution affecting the kinematic solution of the other. Continuous X-ray images of a squat activity for each subject were taken using a flat panel detector [Hitachi, Clavis, Tokyo, Japan; 3 frames/s, image area size 397 (H) $\times$ 298 (V) mm, and 0.20×0.20 mm/pixel resolution]. These images were scaled to 512×512 square pixels for 3D shape registration. A Canny edge detector was used to identify bony contours. At first, bone models were aligned manually by the order of 0.27 mm for in-plane translations, 0.95 mm for out-of-plane translation, and 0.25° for rotation. Next, an automated matching algorithm, based on nonlinear least squares optimization and an image edge-to-model edge distance criteria, was used to align both sets of bone models to 22 X-ray images for each knee (Fig. 1a). The computation time for the matching algorithm was 10 to 20 s per model (Dell precision 650, Intel Xeon processor, 2.40 GHz, 1.00 GB RAM, under the Windows XP Professional edition). Differences were quantified in terms of RMS errors ($\sqrt{bias^2 + variance}$), where bias is the mean difference and variance is the square root of the standard deviation of the differences. Student's t-test ($p < 0.05$) was used to determine if the RMS errors were significantly different from zero.

Experiment 3

The original full-length CT bone models were registered to 22 X-ray images for each knee as previously described. Synthetic X-ray images (Fig. 1b) were then created by ray tracing (Rhinoceros and Flamingo, Robert McNeel & Associates, Seattle, WA) the full-length CT bone models in these 3D poses. The same automated matching algorithm was then used to align

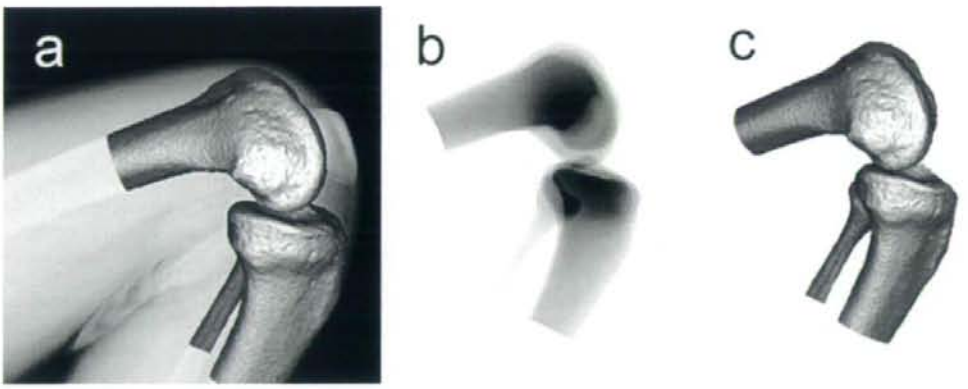

Figure 1. Matching of bone model to X-ray image (a), synthetic image generated using ray tracing (b), and matching of bone model to synthetic image (c).

shortened CT models and the MRI–derived bone models independently to the synthetic images (Fig. 1c).[1,6] RMS errors were used to compare results from the CT and MRI models. Paired t-tests ($p < 0.05$) were used to determine if there were significant differences.

RESULTS

Experiment 1

The differences between CT and MRI model surfaces (mean $\pm$ 1 SD) for the femur and tibia were -0.11 ± 0.81 mm and -0.14 ± 0.67 mm in subject 1, -0.23 ± 0.48 mm and -0.13 ± 0.48 mm in subject 2, and -0.12 ± 0.60 mm and -0.15 ± 0.77 mm in subject 3 (Fig. 2).

Experiment 2

Significant differences were found in three of six parameters (Table 1). RMS differences averaged 1.2 mm for sagittal plane translations, 2.3 mm for mediolateral translations, and 1.7° for all rotations.

Experiment 3

Average RMS errors for tibiofemoral poses were 0.74 mm for sagittal translations, 2.0 mm for mediolateral translations, and 1.4° for all rotations with MRI models. Average RMS errors were 0.53 mm for sagittal translations, 1.6 mm for mediolateral translation, and 0.54° for all rotations with the CT models (Table 2). The total amount of motion observed for each subject is shown in Table 3.

DISCUSSION

Single-plane X-ray imaging and model-based shape matching appear to provide kinematic measurements with sufficient certainty to assess

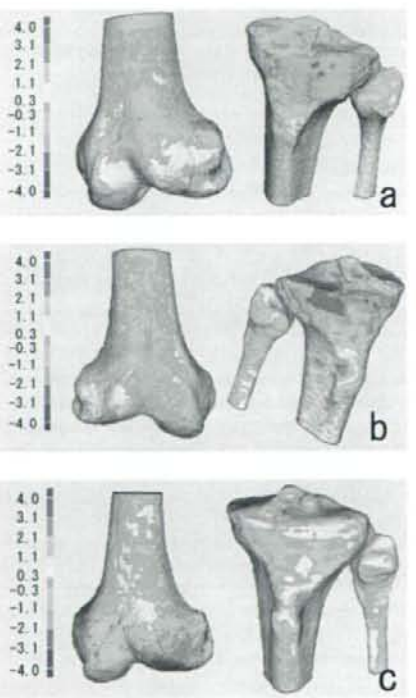

Figure 2. Three-dimensional distance measurement from CT model to MRI model (mm). (a) Subject 1, (b) subject 2, and (c) subject 3.

Table 1. Kinematic Differences When Using MRI– and CT–Derived Bone Models with In Vivo Images (RMS Differences)

Parameter	Subject 1	Subject 2	Subject 3	Average
Anterior–posterior translation (mm)	0.96	0.94	2.28	1.39
Superior–inferior translation (mm)	1.05	1.16	0.84	1.02
Medial–lateral translation (mm)	2.68	1.89	2.32	2.30*
Flexion–extension (°)	1.06	1.17	1.49	1.24*
Internal–external rotation (°)	2.35	0.92	1.54	1.60*
Varus–valgus (°)	1.82	3.42	1.71	2.32

*$p < 0.05$ from 0.

normal and pathological knee motions using either MRI– or CT–derived bone models. Nonetheless, measurement performance with CT–derived bone models was superior to measurements performed with MRI–derived models.

The results with CT–derived models improved significantly from our previous study, where only the exterior bone contours were used for shape matching.[6] Using only exterior contours resulted in average RMS errors of 1.8 mm for sagittal plane translations, 10.6 mm for mediolateral translations, and 1.1° for rotations, compared to 0.53 mm, 1.6 mm, and 0.54°, respectively, for rotations in the present study. Including internal edges for shape registration, specifically the occluded condyles of the femur and tibia and the head of the fibula, significantly improved shape-matching performance.[2] Kinematic measurement performance using single-plane fluoroscopic projections and CT–derived bone models was previously reported.[2] Komistek and colleagues reported measurement precision of 0.45 mm for sagittal plane translation and 0.66° for rotation, both comparable to our current study. Our results are also comparable to measurements using biplane radiographics and CT–derived bone models: RMS error of 0.23 mm for translations and 1.2° for rotation.[7] However, biplane techniques have more uniform errors, whereas single-plane techniques have much higher uncertainties for translations perpendicular to the image plane. No previous study exists to compare for the results with MRI–derived bone models.

Bony contours are less distinct in X-ray projections than the boundaries of metallic implants.[7] Fregly and colleagues showed that biased edge detection can be a primary factor limiting bone-model registration accuracy.[6] Thus, it is critical to adjust X-ray exposure parameters carefully to achieve good contrast around and within the joint. This is especially true for the tibia and fibula, where the fibular head tends to be obscured by the tibia and dense surrounding soft tissue, and the tibial condyles and tubercle are easily overpenetrated by the X-ray beam. Clearly defined, these bony features significantly reduce measurement uncertainty for tibial varus–valgus and internal–external rotation. Use of a high-resolution flat panel detector in the present study permitted adequate definition of these bone features. Image resolution also will affect shape-matching performance and measurement bias. Hardware and software limitations required the flat-panel detector images (1985 × 1490 pixels) to be resampled to 512 × 512 pixels for this analysis. Using higher resolution images, measurement bias could be decreased and measurement precision increased regardless of bone model source.

Table 2. Mean Kinematic Errors When Using MRI– and CT–Derived Bone Models with Synthetic Images (RMS Error)

	Subject 1		Subject 2		Subject 3		Overall		
Tibiofemoral Kinematics	MRI	CT	MRI	CT	MRI	CT	MRI	CT	*t*-Test
Anterior–posterior translation (mm)	0.60	0.44	0.84	0.56	1.19	0.97	0.88	0.66	0.024
Superior–inferior translation (mm)	0.68	0.40	0.46	0.35	0.64	0.44	0.59	0.40	0.057
Medial–lateral translation (mm)	1.54	1.96	2.69	1.34	1.69	1.47	1.97	1.59	0.54
Flexion–extension (°)	0.46	0.27	0.88	0.55	0.98	0.43	0.77	0.42	0.077
Internal–external rotation (°)	2.11	0.43	0.78	0.59	1.41	0.63	1.43	0.55	0.18
Varus–valgus (°)	2.28	0.61	1.34	0.64	1.93	0.67	1.85	0.64	0.049

DOI 10.1002/jor

Table 3. Total Amount of Tibiofemoral Motion for 22 Images in Three Subjects (Maximum Value–Minimum Value)

Tibiofemoral Kinematics	Subject 1	Subject 2	Subject 3	Overall
Anterior–posterior translation (mm)	7.4	11.2	12.2	10.3
Superior–inferior translation (mm)	5.3	3.2	7.2	5.2
Medial–lateral translation (mm)	4.7	6.3	8.4	6.5
Flexion–extension (°)	138.4	115.8	141.5	131.9
Internal–external rotation (°)	32.76	26.48	37.70	32.31
Varus–valgus (°)	6.03	5.54	8.07	6.55

Translations were measured as the femoral coordinate origin moving with respect to the tibial coordinate origin. The amount of motion corresponds to 22 frames of data, which do not necessarily include the entire range of squat motion from full extension to full flexion.

Comparison of the bone models derived from the same subject using CT and MRI showed areas where the surfaces differed by several millimeters (Fig. 2). Several factors probably contributed to these shape differences, which result in different shape-matching performance with the CT– and MRI–derived bone models. First, the fact that different shapes are obtained from the CT and MRI scans introduces bias placing the coordinate systems in the two models. These slight offsets in coordinate system origin and orientation result directly in bias when comparing the measurements from the two models, slightly reducing the ability to isolate differences solely due to bone reconstruction fidelity. Second, bone boundaries identified in CT result directly from X-ray projections, while bone boundaries in MRI result from different physical properties. This consideration is particularly relevant at the distal femur and proximal tibia, where articular regions and ligament insertions present structures with graded properties, where the boundaries are likely to differ between CT and MRI modalities. Thus, we should expect that bone models derived from CT scans will provide superior correspondence when used for shape matching with radiographic projections.

Shape matching with in vivo images showed significant RMS differences comparing kinematics from CT– and MRI–derived bone models (Table 1). When matching the in vivo images with CT–derived bone models, no visible discrepancy was noted between the bone edges in the image and the superimposed edges of the model. With the MRI–derived models, small discrepancies between image and model edges were visible after pose optimization in most cases. Kinematics measured with synthetic X-ray projections uniformly showed less bias and better precision when CT–derived bone models were used (Table 2). Because the synthetic images were created using the CT–derived models, the accuracy and precision figures represent an absolute best-case measurement performance for similar projection geometries using the nonlinear least squares optimization method. The RMS errors figures with the MRI–derived models represent the lower boundary of measurement error one might expect using models based on different physical properties.

When using MRI to create bone models, each MRI scanner will perform differently. For this study, a 0.3 T scanner was used with a gradient echo sequence, and this provided images with sufficient bone/soft tissue contrast to identify the bone boundaries. The gradient echo sequence was used to achieve good resolution for bone segmentation,[4] but spin echo sequences are better for spatial distortion if the contrast is sufficient to detect bone boundaries.[18] Distortion increases with higher magnetic fields.[10,11,17] Higher magnetic fields increase signal intensity for better tissue resolution, but chemical shift and susceptibility artifacts also contribute to geometric distortion. Smaller magnetic fields permit narrower signal bandwidths and consequent reductions in noise. Magnetic field inhomogeneity is another source of geometric distortion that decreases with decreasing magnetic field strength.

Magnetic field inhomogeneity depends on more than field strength, being a function of materials and their spatial distribution within the object being scanned. In biological tissues, MRI signals are generated by hydrogen atoms, with water and fat content accounting for the majority of the signal. All soft tissues and cancellous bone contain a large fraction of water, so the magnetic susceptibility can be approximated by that of water. In contrast, cortical bone and air do not generate significant MRI signals. Nevertheless cortical bone can distort magnetic fields in nearby tissues that do generate

MRI signals, thereby resulting in geometric distortion near these interfaces.[16]

Finally, small motions of the patient during scanning can degrade boundary resolution and spatial integrity of the resulting models. This is of particular concern when sequences requiring long scan times are used, when the anatomy of interest is affected by normal breathing movements, and when immobilization of the area is not easily accomplished. We took great care to reduce motion artifacts while subjects were being scanned, yet it is likely small motion artifacts affected the shape of the MRI-derived bone models. Investigators should attend carefully to positioning and immobilization of subjects to produce high fidelity bone models with MRI.

Useful kinematic measurements can be obtained from single-plane fluoroscopy and shape matching using bone models derived from CT or MRI. Because the fidelity of MRI-derived bone models is degraded by a variety of technical and practical factors, shape-matching results typically will be inferior to those obtained with CT-derived bone models. However, many clinical and research situations exist in which bone model creation using MRI is highly desirable. In these cases, investigators should maintain keen awareness of the factors influencing the fidelity of bone models, and they should incorporate these technical limitations into the interpretation of their findings. Carefully done and cautiously interpreted, we should be able to expand the range of useful kinematic observations using MRI-derived bone models.

ACKNOWLEDGMENTS

No external funds were received in support of this research.

REFERENCES

1. Banks SA, Hodge WA. 1996. Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. IEEE Trans Biomed Eng 43: 638–649.
2. Komistek RD, Dennis DA, Mahfouz M. 2003. In vivo fluoroscopic analysis of the normal human knee. Clin Orthop 69–81.
3. Yamazaki T, Watanabe T, Nakajima Y, et al. 2004. Improvement of depth position in 2-D/3-D registration of knee implants using single-plane fluoroscopy. IEEE Trans Med Imaging 23:602–612.
4. DeFrate LE, Sun H, Gill TJ, et al. 2004. In vivo tibiofemoral contact analysis using 3D MRI-based knee models. J Biomech 37:1499–1504.
5. Asano T, Akagi M, Tanaka K, et al. 2001. In vivo three-dimensional knee kinematics using a biplanar image-matching technique. Clin Orthop 157–166.
6. Fregly BJ, Rahman HA, Banks SA. 2005. Theoretical accuracy of model-based shape matching for measuring natural knee kinematics with single-plane fluoroscopy. J Biomech Eng 127:692–699.
7. You BM, Siy P, Anderst W, et al. 2001. In vivo measurement of 3-D skeletal kinematics from sequences of biplane radiographs: application to knee kinematics. IEEE Trans Med Imaging 20:514–525.
8. Li G, Wuerz TH, DeFrate LE. 2004. Feasibility of using orthogonal fluoroscopic images to measure in vivo joint kinematics. J Biomech Eng 126:314–318.
9. Doran SJ, Charles-Edwards L, Reinsberg SA, et al. 2005. A complete distortion correction for MR images: I. Gradient warp correction. Phys Med Biol 50:1343–1361.
10. Hawnaur JM, Isherwood I. 2000. MRI at midfield strength. In: Young IR editor. Methods in biomedical magnetic resonance imaging and spectroscopy. Chichester, UK; Wiley: p 39–47.
11. Hill DL, Maurer CR Jr, Studholme C, et al. 1998. Correcting scaling errors in tomographic images using a nine degree of freedom registration algorithm. J Comput Assist Tomogr 22:317–323.
12. Kaufman L, Kramer D, Carlson J, et al. 2000. Low-field whole body systems. In: Young IR editor. Methods in biomedical magnetic resonance imaging and spectroscopy. Chichester, UK; Wiley: p 30–39.
13. Langlois S, Desvignes M, Constans JM, et al. 1999. MRI geometric distortion: a simple approach to correcting the effects of non-linear gradient fields. J Magn Reson Imaging 9:821–831.
14. Maurer CR Jr, Aboutanos GB, Dawant BM, et al. 1996. Effect of geometrical distortion correction in MR on image registration accuracy. J Comput Assist Tomogr 20:666–679.
15. Moore CS, Liney GP, Beavis AW. 2004. Quality assurance of registration of CT and MRI data sets for treatment planning of radiotherapy for head and neck cancers. J Appl Clin Med Phys 5:25–35.
16. Sumanaweera T, Glover G, Song S, et al. 1994. Quantifying MRI geometric distortion in tissue. Magn Reson Med 31:40–47.
17. Vlaardingerbroek MT, Boer JAd. 1999. Magnetic resonance imaging. 2nd ed. New York: Springer; p 185–187.
18. Young IR editor. 2000. Biomedical magnetic resonance imaging and spectroscopy. Chichester, UK: Wiley.
19. Eckhoff DG, Bach JM, Spitzer VM, et al. 2005. Three-dimensional mechanics, kinematics, and morphology of the knee viewed in virtual reality. J Bone Joint Surg [Am] 87: (Suppl 2):71–80.
20. Roos PJ, Neu CP, Hull ML, et al. 2005. A new tibial coordinate system improves the precision of anterior-posterior knee laxity measurements: a cadaveric study using Roentgen stereophotogrammetric analysis. J Orthop Res 23:327–333.

J Orthop Sci (2007) 12:214–218
DOI 10.1007/s00776-007-1112-7

JOURNAL OF
ORTHOPAEDIC SCIENCE

Original article

Femoral shaft bowing influences the correction angle for high tibial osteotomy

RYUJI NAGAMINE[1], SANSHIRO INOUE[2], HIROMASA MIURA[2], SHUICHI MATSUDA[2], and YUKIHIDE IWAMOTO[2]

[1]Department of Orthopaedic Surgery, Yoshizuka Hayashi Hospital, 7-6-29 Yoshizuka, Hakata-ku, Fukuoka, Fukuoka 812-0041, Japan
[2]Department of Orthopaedic Surgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

Abstract

Background. The optimal femorotibial angle (FTA) after high tibial osteotomy (HTO) is still controversial. Our hypothesis was that FTA itself may not be reliable because FTA cannot represent the accurate alignment of the whole lower extremity.

Methods. Non-weight-bearing radiographs of the lower extremities were taken in 100 Japanese subjects with medial osteoarthritic knees, and seven anatomic parameters were assessed. The correction angle by FTA was calculated so that the postoperative FTA was set at 166° (14° valgus). Another correction angle was calculated so that the mechanical axis passed through the lateral one-fourth of the tibial articular surface after HTO. After the correlation between two correction angles was assessed, influences of anatomic parameters on the discrepancy between two correction angles were assessed.

Results. There was a high correlation between two correction angles ($R^2 = 0.777$, $P < 0.001$). The mechanical axis passed through the lateral one-fourth of the tibial articular surface when the postoperative FTA was set at 166° in 80% of subjects. However, discrepancy between the two correction angles was 3° or larger in 20% of subjects. Femoral shaft bowing and tibial shaft bowing significantly influenced the correction angles. Even though FTA was the same, the femoral head shifted medially in cases with lateral bowing of the femoral shaft, and the correction angle by FTA should be set larger. On the other hand, the correction angle by FTA can be set smaller in knees with medial bowing of the femoral shaft. Tibial shaft bowing also influences the correction angle by FTA.

Conclusions. The correction angle by FTA for HTO should be calculated taking femoral and/or tibial shaft bowing in the frontal plane into account.

Offprint requests to: R. Nagamine
Received: September 4, 2006 / Accepted: January 10, 2007

Introduction

High tibial osteotomy (HTO) is one of the surgical treatments for medial osteoarthritis of the knee. The knee alignment before and after HTO is mainly evaluated by the femorotibial angle (FTA).[1–13] Long-term follow up studies had shown the optimal postoperative FTA to be between 164° and 170° (16° valgus and 10° valgus) in Japan,[1,2] and this angle was between 164° and 175° (16° valgus and 5° valgus) in Western countries.[3–6] Some studies had shown no relationship between the postoperative FTA and clinical outcome.[7–9] The optimal FTA after HTO is still obscure. It has never been analyzed why the optimal FTA after HTO was different among the studies. Our hypothesis was that FTA itself may not be reliable, because FTA does not represent the accurate alignment of the whole lower extremity, and anatomic variations of the femur and the tibia in the frontal plane are not taken into account. The mechanical axis can be used to calculate the other correction angle for HTO.[14] The correction angle by the mechanical axis is not influenced by anatomic variations; therefore, it may be more reliable than the correction angle by FTA. The purpose of this study is to clarify whether FTA is reliable for evaluation of the alignment of the lower extremity in all subjects and whether anatomic variations in the frontal plane influence the correction angle by FTA.

Materials and methods

This study consisted of 100 lower extremities in 100 Japanese subjects with medial osteoarthritic knees. Interlocking closed-wedge HTO was performed in all subjects.[14] There were 22 men and 78 women, and their mean age was 68 years old (range, 50–82 years). On radiography in the standing position, the medial femorotibial joint space had disappeared in 25 cases and was

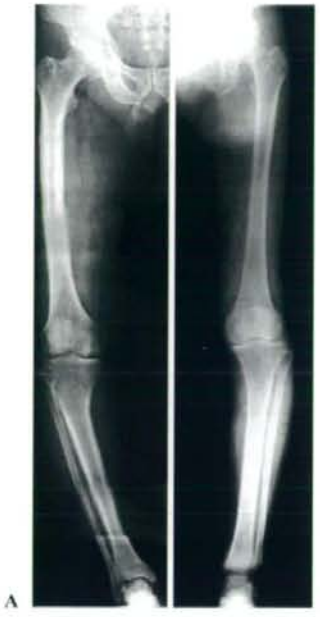

Fig. 1. Orthoradiography was taken in 100 lower extremities of 100 cases in that HTO was performed. Two cases with lateral (**A**) and medial (**B**) bowing of the femoral shafts are shown

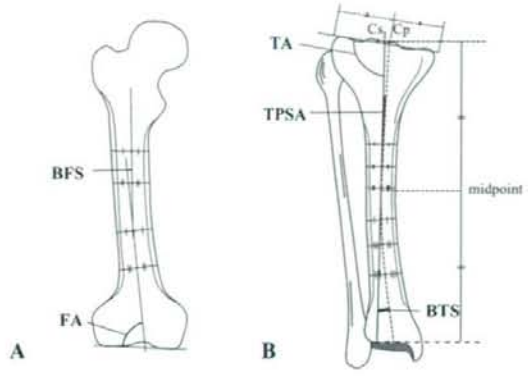

Fig. 2. Schematics of the parameters are shown: *FA*, Femoral angle; *BFS*, lateral bowing angle of the femoral shaft (**A**); *Cp*, central point of the tibial articular surface; *Cs*, central line of the proximal tibial shaft; *midpoint*, point located by bisecting the proximal-to-distal length of the tibia; *TA*, tibial angle; *TPSA*, tibial plateau shift angle; *BTS*, lateral bowing angle of the tibial shaft (**B**)

less than 3 mm in 52 cases. It was slightly decreased in 23 cases. According to the rules of the ethics committee in our hospitals, all patients were informed that data from the study would be submitted for publication, and they gave their consent. Before HTO, all patients received the standardized non-weight-bearing radiographs of their lower extremities, known as orthoradiography, which used three successive exposures centered over the hip, knee, and ankle, with tube-to-film distance of 2 m (Fig. 1).[14,15] Care was taken to place the lower extremities in a neutral position so that the patella faced anteriorly. No patient had flexion contracture of more than 10°. No patient had patellar subluxation/dislocation. On orthoradiography, central lines of proximal and distal diaphysis of the femur and the tibia were drawn.[15] Seven anatomic parameters were measured, these being the femoral angle (FA), the lateral bowing angle of the femoral shaft (BFS), the tibial angle (TA), the lateral bowing angle of the tibial shaft (BTS), the tibial plateau shift angle (TPSA), FTA, and the percentile of the mechanical axis on the tibial plateau (%MA) (Fig. 2).[15] FA is the lateral angle between the central line of the distal diaphysis of the femur and the tangent of the distal femoral condyles. BFS is the angle between the central lines of the distal and proximal diaphysis of the femur. TA is the lateral angle between the central line of the proximal diaphysis of the tibia and the line between the medial and lateral tibial plateaus. BTS is the angle between the central lines of the distal and proximal diaphysis of the tibia. TPSA is the angle between the central line of the proximal diaphysis of the tibia and the line from the central point of the tibial plateau (articular surface) to the midpoint of the tibia. TPSA demonstrates the medial deviation of the central point of the tibial articular surface from the central line of the proximal diaphysis of the tibia. FTA is the lateral angle between the central line of the distal diaphysis of the femur and the central line of the proximal diaphysis of the tibia. The mechanical axis was drawn from the central point of the femoral head to the central point of the articular surface of the talus. %MA was 0% when the mechanical axis passed through the medial edge of the tibial plateau and was 100% when the mechanical axis passed through the lateral edge.

When the correction angle by %MA was calculated on the orthoradiograph, postoperative alignment was achieved so that %MA was 75% (lateral one-fourth of the tibial articular surface).[11,14] The method of the calculation is shown in Fig. 3.[14] First, the preoperative mechanical axis is drawn. The broken line shows an osteotomy line parallel to the tibial articular surface (Fig. 3A). The postoperative mechanical axis is drawn so that %MA is 75% (Fig. 3B). A line is drawn from the medial edge of the osteotomy line to the center of the talus, defined as line A. Line B, which has the same length as line A, is drawn from the medial edge of the osteotomy line to the postoperative mechanical axis

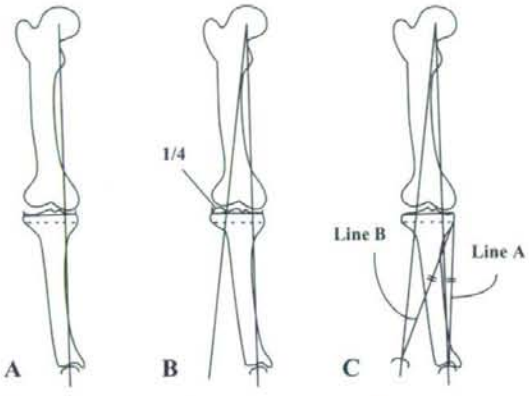

Fig. 3. Schematic of correction angle by %MA (percentile of the mechanical axis on the tibial plateau) is shown. The *broken line* shows the osteotomy line (**A**). The postoperative mechanical axis is drawn so that %MA is 75% (**B**). The angle between *line A* and *line B* shows the correction angle by %MA (**C**)

(Fig. 3C). The angle between line A and line B is to be the correction angle by %MA. While the correction angle by FTA was calculated, the postoperative FTA was set at 166° (14° valgus), because it has been reported that FTA of 164° to 168° (16° to 12° valgus) should be attained to ensure favorable long-term results in HTO in Japan.[1] If the discrepancy between two correction angles was larger or equal to 3°, the correction angles were considered to be unreliable in this study.[1,16]

The correlation between two correction angles was assessed. Next, the correlation between each anatomic parameter and the discrepancy between two correction angles was assessed. Statistical analysis was done using the correlation analysis, with a probability of less than 0.05 being significant.

Results

FA was 81.9° ± 2.1° (mean ± standard deviation) (range, 77° to 87°); BFS was 2.2° ± 3.2° (–3.5° to 16°); TA was 97.2° ± 3.4° (89° to 106°); BTS was –0.1° ± 2.0° (–6° to 6°); and TPSA was 2.1° ± 1.5° (0° to 7°). The preoperative FTA was 180.8° (0.8° varus) ± 4.8° (171° to 196°) (9° valgus to 16° varus) and the preoperative %MA was 19.6% ± 15.6% (–25.4% to 47%). The correction angle by FTA and that by %MA were calculated, and they were 14.8° ± 4.8° (5° to 30°) and 15.1° ± 4.7° (7.5° to 28°), respectively. The results showed that the characteristics of Japanese patients with medial osteoarthritic knees were the proximal tibia vara with medial shift of the tibial articular surface and lateral bowing of the femoral shaft.

There was a high correlation between two correction angles ($R^2 = 0.777$, $P < 0.001$). Both the mean correction angle by FTA and that by %MA were about 15°. The discrepancy between two correction angles was less than 3° in 80 cases. Therefore, the mechanical axis passed through the lateral one-fourth of the tibial articular surface when the postoperative FTA was set at 166° (14° valgus) in 80% of cases. However, the discrepancy between two correction angles was 3° or larger in 20% of the cases. Three parameters had significant correlation with the discrepancy: BFS ($P < 0.001$), BTS ($P < 0.01$), and TA ($P < 0.05$). The correction angle by %MA was greater than that by FTA by 3° or larger in 14 cases. The femur had lateral bowing of the femoral shaft in such cases. The correction angle by %MA was smaller than that by FTA by 3° or larger in six cases. The femur had medial bowing in such cases. Figure 1 shows two cases with lateral and medial bowing of the femoral shaft. In the case with lateral bowing of the femoral shaft (Fig. 1A), the correction angle by %MA was 26° and that by FTA was 23°, whereas in the case with medial bowing of the femoral shaft (Fig. 1B), the correction angles were 12° and 16°, respectively. These results clearly showed that the anatomic variations in the frontal plane influence the correction angle by FTA, and that FTA may not demonstrate the accurate alignment for the whole lower extremity in 20% of Japanese subjects.

Discussion

The correction angle is one of the most important factors that influence the outcome of HTO. The correction angle has been calculated only by FTA in many studies, and the degrees of anatomic variations of the femur and the tibia in the frontal plane have not been taken into consideration as one of factors that influenced the outcome. The results of this study showed that the correction angle by FTA may not be reliable in cases with bowing of the femoral and/or tibial shaft. BFS was the most important variation that influenced the correction angle. Even though FTA is the same, the mechanical axis passes through the more medial side of the knee in cases with lateral bowing of the femoral shaft because the femoral head shifts medially (Fig. 4). In such cases, the correction angle by FTA should be set larger. On the other hand, the postoperative FTA can be set more varus in knees with medial bowing of the femoral shaft. The correction angle by FTA was smaller than that by %MA in the case with lateral bowing of the femoral shaft (Fig. 1A). The correction angle by FTA was larger than that by %MA in the case with medial bowing of the femoral shaft (Fig.

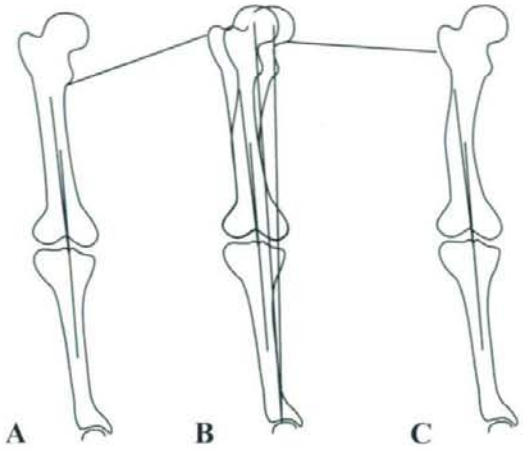

Fig. 4. Bowing of the femoral shaft that influences the correction angle is shown. The case in **A** has a straight femoral shaft. The case in **C** has lateral bowing of the femoral shaft. In **B**, **A** and **C** are superimposed

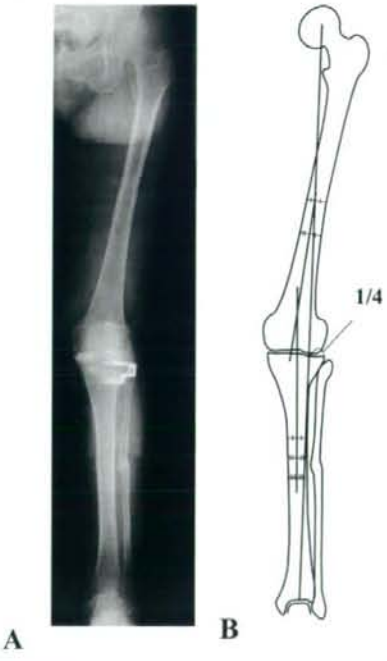

Fig. 5. Orthoradiography (**A**) and schematic (**B**) of the lower extremity of the case of Fig. 1B at 8 years after HTO. Even though FTA is 170°, the mechanical axis passes though the lateral one-fourth of the tibial articular surface

1B). Figure 5 shows the orthoradiography of the case with medial bowing of the femoral shaft (Fig. 1B) at 8 years after HTO. Even though FTA was 170°, %MA was 75%, and the knee function score of the Japan Orthopaedic Association was 92.[1,2]

The results showed that the mechanical axis will pass through the lateral one-fourth of the tibial articular surface when FTA is set at 166° (14° valgus) in 80% of cases in Japan. Fourteen degrees is larger than the optimal FTA that has been reported from Canada and the United States.[3,4] It has been reported that FA was the same between Japanese subjects and subjects in Western countries.[15] These results showed that Japanese subjects in this study might have varus deformity of the femoral shaft and/or the tibial shaft in addition to the proximal tibia vara. The anatomic characteristics of the lower extremities are different among the countries (races).[15] When clinical results of HTO are evaluated, not only FTA but also the accurate alignment of the whole lower extremity should be assessed, because patients may have some degrees of anatomic variation of their femoral and/or tibial shafts in the frontal plane. When the alignment of the whole lower extremity was compared among different countries, there are several problems. First, the definition of FTA is different among the countries.[17] Therefore, the value of FTA was expressed in two ways in this study. The hip-knee-ankle angle is popular to express the alignment of the whole lower extremity in the United States.[16,18] Studies among Japan and other countries are necessary to obtain optimal correction angle for HTO. In those studies, the accurate alignment of the whole lower extremities should be evaluated using unified parameters.

In conclusion, the correction angle by FTA for HTO should be calculated taking femoral and/or tibial shaft bowing in the frontal plane into account.

Acknowledgments. The authors thank Miss Tomoko Yoshida and Miss Chiemi Kadota for their technical assistance.

References

1. Yasuda K, Majima T, Tsuchida T, Kaneda K. A ten- to 15-year follow-up observation of high tibial osteotomy in medial compartment osteoarthrosis. Clin Orthop 1992;282:186–95.
2. Choi HR, Hasegawa Y, Kondo S, Shimizu T, Ida K, Iwata H. High tibial osteotomy for varus gonarthrosis: a 10- to 24-year follow-up study. J Orthop Sci 2001;6:493–7.
3. Rudan JF, Simurda MA. High tibial osteotomy. A prospective clinical and roentgenographic review. Clin Orthop 1990;255: 251–6.
4. Coventry MB, Ilstrup DM, Wallrichs SL. Proximal tibial osteotomy: a critical long-term study of eighty-seven cases. J Bone Joint Surg [Am] 1993;75:196–201.
5. Aglietti P, Buzzi R, Vena LM, Baldini A, Mondaini A. High tibial valgus osteotomy for medial gonarthrosis: a 10- to 21-year study. J Knee Surg 2003;16:21–6.

6. Sprenger TR, Doerzbacher JF. Tibial osteotomy for the treatment of varus gonarhthrosis. Survival and failure analysis to twenty-two years. J Bone Joint Surg [Am] 2003;85:469–74.
7. Huang TL, Tseng KF, Chen WM, Lin RM, Wu JJ, Chen TH. Preoperative tibiofemoral angle predicts survival of proximal tibial osteotomy. Clin Orthop 2005;432;188–95.
8. Sundaram NA, Hallett JP, Sullivan MF. Dome osteotomy of the tibia for osteoarthritis of the knee. J Bone Joint Surg [Br] 1986;68: 782–6.
9. Naudie D, Bourne RB, Rorabeck CH, Bourne TJ. Survivorship of the high tibial valgus osteotomy. A 10- to 22- year follow up study. Clin Orthop 1999;367:18–27.
10. Dearborn JT, Eakin CL, Skinner HB. Medial compartment arthrosis of the knee. Am J Orthop 1996;25:18–26.
11. Dugdale TW, Noyes FR, Styer D. Pre-operative planning for high tibial osteotomy. Clin Orthop 1992;274:248–64.
12. Madan S, Ranjith RK, Fiddian NJ. Intermediate follow-up of high tibial osteotomy. A comparison of two techniques. Bull Hosp Jt Dis 2002–2003;61:11–6.
13. Nagel A, Insall JN, Scuderi GR. Proximal tibial osteotomy: a subjective outcome study. J Bone Joint Surg [Am] 1996;78: 1353–8.
14. Ogata K, Yoshii I, Kawamura H, Miura H, Arizono T, Sugioka Y. Standing radiographs cannot determine the correction in high tibial osteotomy. J Bone Joint Surg [Br] 1991;73:927–31.
15. Nagamine R, Miura H, Bravo CV, Urabe K, Matsuda S, Miyanishi K, et al. Anatomic variations should be considered in total knee arthroplasty. J Orthop Sci 2000;5:232–7.
16. Magyar G, Ahl TL, Vibe P, Toksviq-Larsen S, Lindstrand A. Open-wedge osteotomy by hemicallotasis or the closed-wedge technique for osteoarthritis of the knee. A randomized study of 50 operations. J Bone Joint Surg [Br] 1999;81:444–8.
17. Moreland JR, Bassett LW, Hanker GJ. Radiographic analysis of the axial alignment of the lower extremity. J Bone Joint Surg [Am] 1987;69:745–9.
18. Stukenborg-Colsman C, Wirth CJ, Lazovic D, Wefer A. High tibial osteotomy versus unicompartmental joint replacement in unicompartmental knee joint osteoarthritis: 7–10-year follow-up prospective randomized study. Knee 2001;8:187–94.