The Impact of Leg Length Discrepancy on Clinical Outcome of Total Hip Arthroplasty: Comparison of Four Measurement Methods

Matej Keršič, MD, Drago Dolinar, MD, PhD, Vane Antolič, MD, PhD, Blaž Mavčič, MD, PhD
Department of Orthopaedic Surgery, University Medical Centre Ljubljana, Ljubljana, Slovenia

ABSTRACT

In a single-surgeon series of 119 patients with unilateral primary un cemented total hip arthroplasty, four leg-length discrepancy measurement methods (absolute, relative, trochanteric, standardized-trochanteric) were analyzed for their impact on WOMAC score, Oxford Hip Score and self-perceived leg-length discrepancy. After adjustment for age, gender and BMI, postoperative WOMAC scores correlated only with clinical absolute measurements of leg elongation (P = 0.05). Self-perceived leg-length discrepancy corresponded best to the clinically measured relative leg-length discrepancy (11 mm perceived vs. 7 mm unperceived; P = 0.04) while there was no significant correspondence with radiographic measurements or leg elongation magnitudes. Within the <10 mm range of mean postoperative leg-length discrepancy in the studied series, its impact on the overall clinical satisfaction was detectable but not considerable.

Keywords: arthroplasty hip leg length inequality patient outcome assessment

Leg-length discrepancy (LLD) is a well recognized complication of total hip arthroplasty (THA) [1–8]. It is generally accepted that excessive LLD magnitudes over 1.5 cm can cause lower back pain [9,10], gait disorders [11–13] and general dissatisfaction [14–16]. However, the impact of LLD magnitudes below 1.5 cm on the overall patient satisfaction with THA is less clear. While some authors report even small differences in leg length after THA have a negative influence on limping and pain [16–20], other recent studies show no significant correlation between radiographic LLD measurements and patient-reported satisfaction [21–23].

One of the reasons for conflicting reports on LLD in patients undergoing THA is that four different LLD measurement methods (two clinical methods: absolute and relative and two radiological methods: trochanteric and standardized-trochanteric) can be used in clinical practice [1,24,25]. The absolute LLD is the difference of clinically measured distances from anterior superior iliac spine to the inner malleolus on each side. With the relative method, a tape is used to clinically measure LLD as a difference of distances from umbilicus to each inner malleolus and thus this method also accounts for pelvic obliquity, contractures in lumbar spine, hip and knee. The trochanteric method involves radiographic comparison of the vertical distance of lesser trochanters from the interteardrop line. With the standardized-trochanteric method the vertical distance of lesser troCHANTERS FROM THE INTERTEARDROP LINE IS MEASURED IN A SEPARATE COORDINATE SYSTEM FOR EACH FEMUR IN ORDER TO AVOID MEASUREMENT ERRORS DUE TO PELVIC AND/OR FEMORAL INCLINATION, AS SHOWN IN FIG. 1 [25]. MEASUREMENT ERRORS, VALIDITY, INTRA-OBSERVER AND INTER-OBSERVER RELIABILITY OF VARIOUS LLD MEASUREMENT METHODS HAVE BEEN THOROUGHLY INVESTIGATED IN THE PAST [26–29]. MOST PUBLISHED REPORTS ON LLD AFTER THA ARE BASED ON RADIographic ANALYSES BECAUSE THEY ARE THOUGHT TO ENABLE MORE PRECISE MEASUREMENTS THAN CLINICAL METHODS. HOWEVER, VARIABILITY IN PATIENT POSITIONING AND FEMUR INCLINATION/rotation CAN ALSO CAUSE SIGNIFICANT MEASUREMENT ERRORS OF RADIOgraphic METHODS [30]. IN ADDITION, THE TROCHANTERIC AND THE STANDARDIZED-TROCHANTERIC METHOD ONLY TAKE INTO ACCOUNT LEG-LENGTH INEQUALITY IN THE PROXIMAL FEMUR AND DISREGARD POSSIBLE PREOPERATIVE SUBTROCHANTERIC CAUSES OF LLD [31,32]. THE QUESTION THEREFORE REMAINS TO BE ANSWERED WHICH OF THE FOUR LLD MEASUREMENT METHODS BETTER PREDICTS SUBJECTIVELY PERCEIVED LLD IN THE POSTOPERATIVE SETTING. FURTHERMORE, FROM REPORTS IN THE LITERATURE IT IS NOT CLEAR WHETHER POSTOPERATIVE FUNCTIONAL PROBLEMS DERIVE FROM THE LEG ELONGATION PER SE OR FROM THE POSTOPERATIVE LLD [22]. AFTER THA, HIP PAIN AND CONTRACTURES ARE GREATLY REDUCED, HIP ROTATION CENTER IS DISPLACED, OFFSET AND LEG-LENGTH CHANGE [33,34] AND IT IS DIFFICULT TO ISOLATE THE INDEPENDENT INFLUENCE OF POSTOPERATIVE LLD ON THE OVERALL PATIENT SATISFACTION.

The aim of this paper was to evaluate the differences between four LLD measurement methods (absolute, relative, trochanteric, standardized-trochanteric) when measuring preoperative LLD, postoperative LLD and leg elongation in a single group of patients undergoing THA. Further we wanted to find out which of the four LLD measurement methods predicted best the patient-perceived postoperative LLD and the overall clinical outcome of THA. Finally we...
ascertained the impact of leg elongation magnitude on the patient-perceived postoperative LLD and the overall clinical outcome of THA.

Materials and Methods

Patients

The study included patients with unilateral unementation primary THA (implants Zweymüller-EndoPlus, Wright-Profemur2 or EcoFit-Implantcast) who were operated by a single experienced surgeon (D.D.) at our institution in the period between January 1, 2004, and June 30, 2011, without postoperative luxations, infections or revision operations. Out of 630 primary and revision THA performed by this surgeon in the selected time period, 119 patients met the inclusion criteria. The data were analyzed retrospectively at least 1 year after the surgery. The indications for THA included idiopathic osteoarthritis in 98 cases, avascular necrosis of the femoral head in 10 cases, hip dysplasia in 8 cases, posttraumatic osteoarthritis in 2 cases and one case of osteoarthritis due to the Legg-Calve-Perthes disease in childhood. All patients were operated through the lateral translgluteal approach in the supine position. At the time of operation the surgeon was acquainted with the values of preoperative clinical LLD measurements. His intraoperative aim was to lengthen the operated leg according to the preoperative absolute LLD and after trial THA reduction he would manually assess the final leg lengths by comparing malleolus and patellae in the supine position. The wound was drained for 24 hours, the single dose cephalozin preoperative antibiotic prophylaxis and the LMWH postoperative antithrombotic prophylaxis were given. Patients were mobilized with crutches on the postoperative day 1, full weight-bearing was allowed after 6 weeks. All patients were followed up at least 1 year postoperatively. None of the patients exhibited postoperative Trendelenburg gait at 1 year follow-up.

The study protocol was reviewed and approved by the National Medical Ethics Committee of the Republic of Slovenia on September 28, 2010. All participants signed an informed consent form for participation in the study.

Leg-Length Discrepancy Measurements

Leg-length discrepancy (LLD) was measured in each patient preoperatively and postoperatively with four different measurement methods (absolute, relative, trochanteric, standardized-trochanteric). With the absolute method, a tape was used to measure LLD as the difference in lengths from anterior spina to inner malleolus on both sides. With the relative method, a tape was used to measure LLD as the difference in lengths from umbilicus to each inner malleolus [25]. The trochanteric method involved measuring LLD as the difference in vertical distance of lesser trochanters from the interteardrop line. The standardized-trochanteric method involved measurements of the shift in hip rotation center (preoperative vs. postoperative) according to two separate coordinate systems (pelvic and femoral), thus avoiding the influence of pelvic and femoral inclination on the radiographs [25]. The standardized-trochanteric LLD was defined as the difference in vertical distance of lesser trochanters from the interteardrop line whereby the vertical reference line was specific for each femur and parallel to its longitudinal axis.

Measurements were performed by two independent investigators and the operating surgeon was excluded from the measurement process to ensure objectivity. One investigator performed clinical measurements (absolute, relative), the other investigator performed radiographic measurements (trochanteric, standardized-trochanteric) and they were blinded for each other’s results. Clinical tape measurements were performed with the nearest reading of 5 mm and radiographic measurements were performed with the nearest reading of 1 mm. Leg elongation at THA was calculated by subtracting the preoperative LLD from the postoperative LLD for each of the four measurement methods.

Preoperative Clinical Assessment and Evaluation of the THA Clinical Outcome

Preoperative clinical assessment was performed with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC score), the officially evaluated translation of the WOMAC 3.1-VAS questionnaire in the Slovenian language (Prof. Nicholas Bellamy) [35]. Clinical status of patients after THA and self-reported problems with postoperative LLD were assessed with a custom-made questionnaire, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC score) [35] and the Oxford Hip Score (OHS) [36].

The custom-made questionnaire included the following questions: 1). Did THA improve the quality of Your life? 2). Do you feel that after THA Your legs are equally long? Patients who answered the last question with “NO” were asked additional two questions: 3). Do You use a shoe raise? 4). Do You have more problems with LLD than prior to THA? Patients indicated either “YES” or “NO” for each question.

The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC score) is a patient-reported outcome measure developed to assess pain, stiffness, and physical function in patients with hip and/or knee osteoarthritis. It consists of 24 items divided into 3 subscales: pain (5 items): during walking, using stairs, in bed, sitting or lying, and standing; stiffness (2 items): after first waking and later in the day; physical function (17 items): stair use, rising from sitting, standing, bending, walking, getting in/out of a car, shopping, putting on/taking off socks, rising from bed, lying in bed, getting in/out of bath, sitting, getting on/off the toilet, heavy household duties, light household duties. We used the Visual Analog Scale version where the patient answered each question by marking a cross on a 100 mm line to indicate the present clinical status. A ruler was then used to measure the distance (in mm) from the left end marker to the patient’s mark. For each item, the possible range of scores is 0 mm–100 mm. Items are summed for each subscale and a total WOMAC score (0 mm–2400 mm) is computed by summing the items for all three subscasles. Higher WOMAC scores indicate worse pain, stiffness, and functional limitations [35].

The Oxford Hip Score (OHS) is a patient-reported outcome measure developed to assess functional ability and pain from the patient’s perspective. It is a joint-specific questionnaire developed and validated for use in patients undergoing THA. The OHS consists of 12 questions about pain and physical limitations experienced over the past 4 weeks because of the hip. Each question has five response categories, giving a score between 1 and 5 (no disability to high disability). The total sum of scores for the 12 items gives the final score between 12 and 60 whereby higher score indicates greater level of functional disability [36].

Statistical Methods

Differences between means were evaluated with the two-tailed Mann–Whitney U test for unpaired samples. The difference in proportions between groups was tested with the Fisher’s exact test. The impact of LLD and leg elongation on clinical outcome scores was evaluated with linear regression models by analyzing each of the three different dependent variables (improvement of WOMAC score, postoperative WOMAC score and OHS) with four predictor variables (postoperative LLD or leg elongation magnitude for each of the four measurement methods, age, gender and BMI). Statistical significance was set at $P \leq 0.05$.

To determine the required sample size, power analysis was performed a priori both for comparison of means with the two-tailed Mann–Whitney U test and for the multivariate linear regression models. The effect size for comparison of means with the two-tailed
Mann–Whitney U test was computed from the minimal clinically-detectable leg-length discrepancy (≥5 mm) and from the standard deviation of leg length discrepancy (9 mm) in THA population of our previous pilot study [37]. For α = 0.05 and power (1−β) = 0.80, the required minimal total sample size for comparison of means was 78 subjects. The effect size for linear regression models (fixed model, 1 tested predictor variable + 3 predictor covariates) was computed from the effect size $f^2 = 0.15$ that corresponds to the expected medium effect size [38]. For α = 0.05 and power (1−β) = 0.80, the required minimal total sample size for linear regression models was 55 subjects.

In all statistical calculations preoperative/postoperative LLD measurements were used as positive values regardless of whether the operated leg was shorter or longer from the unoperated leg. Leg elongation magnitudes, however, were included in calculations as either negative or positive values.

Statistical analyses were performed with SPSS 17.0 for Windows (SPSS Inc, Chicago, IL), Microsoft Office Excel 2010 (Microsoft Inc, Redmond, WA), and GPower 3.1.5 (Faul et al.; Düsseldorf, Germany) [38].

Results

Preoperatively the mean LLD of the analyzed 119 patients was $9 \pm 8$ mm measured with the absolute method, $14 \pm 12$ mm with the relative method, $7 \pm 7$ mm with the trochanteric method and $7 \pm 7$ mm with the standardized-trochanteric method. In (62/119) 52% of cases the leg to be operated was shorter, in (12/119) 10% the leg to be operated was longer and in the remaining cases the trochanteric preoperative LLD was lower than 5 mm.

The mean elongation of the operated leg at THA was $8 \pm 9$ mm measured with the absolute method, $13 \pm 13$ mm with the relative method, $10 \pm 9$ mm with the trochanteric method and $9 \pm 8$ mm with the standardized-trochanteric method. Postoperatively the mean LLD of the analyzed 119 patients was $6 \pm 7$ mm measured with the absolute method, $10 \pm 9$ mm with the relative method, $8 \pm 6$ mm with the trochanteric method and $9 \pm 8$ mm with the standardized-trochanteric method. After THA, in (18/119) 15% of cases the operated leg was shorter, in (58/119) 49% the operated leg was longer and in the remaining cases the trochanteric postoperative LLD was lower than 5 mm.

Preoperative WOMAC score was available for 34 patients, their mean value was $1891 \pm 404$ mm, preoperative leg length discrepancy was perceived by (22/34) 65% of these patients. Out of 119 patients who were sent the postoperative questionnaires, 107 patients filled-out the custom-made questionnaire on possible problems with postoperative LLD (Table 1). 93 patients reported the postoperative WOMAC score and 106 patients reported the postoperative Oxford Hip Score. The mean postoperative WOMAC score was $312 \pm 395$ mm and the mean postoperative OHS value was $20 \pm 9$ points.

The 107 patients who postoperatively filled-out custom-made questionnaires were further stratified according to postoperatively perceived LLD (n = 44 patients) or unperceived LLD (n = 63 patients). Between the two groups there was no statistically significant difference in the mean age at operation (60.8 years vs. 60.9 years; $P = 0.88$), the mean age at follow-up (62.8 years vs. 62.9 years; $P = 0.93$), the proportion of female patients ($52 \% \text{ vs.} 51 \%; P = 0.99$), the mean body mass index ($26.9 \text{ kg/m}^2 \text{ vs.} 28.5 \text{ kg/m}^2; P = 0.13$) or the mean preoperative WOMAC score ($1771 \text{ mm vs.} 1986 \text{ mm}; P = 0.08$). However, we found significantly worse clinical outcomes in the group of perceived LLD both for postoperative WOMAC score ($462 \text{ mm vs.} 217 \text{ mm}; P = 0.01$) and OHS score ($23 \text{ vs.} 17; P < 0.01$).

To exclude the confounding factors (age/gender/BMI), the impact of postoperative LLD on clinical scores was analyzed with linear regression for four different LLD measurement methods and stratified according to subjectively perceived or unperceived postoperative LLD (Table 2). Furthermore, the impact of leg elongation at THA on clinical scores was analyzed with linear regression for four different LLD measurement methods and stratified according to subjectively perceived or unperceived postoperative LLD (Table 3).

Discussion

The presented results demonstrate how the differences between the four LLD measurement methods (absolute, relative, trochanteric, standardized-trochanteric) depend on the population studied. In the preoperative setting (hip contracture present) relative LLD measurements on average showed 5–8 mm larger values in comparison with radiographic methods (trochanteric and standardized-trochanteric). Postoperatively when hip contracture was released, the mean LLD measured with relative, trochanteric or standardized-trochanteric method differed only by 1–2 mm. The mean absolute LLD measurement was consistently lower from the mean relative measurement by ≥4 mm, both in preoperative LLD, postoperative LLD and at leg elongation. Furthermore, our findings emphasize the difference between subjectively perceived and objectively measured postoperative LLD. Subjective feeling of postoperative leg inequality corresponded only with relative LLD measurements while leg elongation magnitudes did not significantly differ between the groups with perceived and unperceived LLD. Standardized-trochanteric radiographic method did not show any significant advantage over trochanteric radiographic method in terms of predicting subjectively perceived LLD or overall clinical outcome. After adjustment for age, gender and BMI, linear regression models showed no statistically significant impact of postoperative LLD or leg elongation on clinical scores, except for small correlation between clinically measured absolute leg elongation and postoperative WOMAC score.

### Table 1

<table>
<thead>
<tr>
<th>1. Did THA improve the quality of Your life?</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>103 (96%)</td>
<td>4 (4%)</td>
<td></td>
</tr>
<tr>
<td>2. Do you feel that after THA Your legs are equally long?</td>
<td>63 (59%)</td>
<td>41 (41%)</td>
</tr>
<tr>
<td>Subjects who answered Question No. 2, with “NO” (N = 44 patients) further replied to the following questions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Do You use a shoe raise?</td>
<td>24 (55%)</td>
<td>20 (45%)</td>
</tr>
<tr>
<td>4. Do You have more problems with LLD than prior to THA?</td>
<td>22 (50%)</td>
<td>22 (50%)</td>
</tr>
</tbody>
</table>
Clinical results and magnitudes of preoperative/postoperative LLD measurements in this patient series are comparable to the series of other authors [16]. In this study, the use of a shoe raise for treatment of postoperative LLD was common (55% of the LLD-perceived group used a shoe raise), which supports the conclusions of other authors [39]. While we found the prevalence of patient-perceived LLD after THA was 41% (Table 1), a considerable proportion of these patients (52%) actually had trochanteric LLD ≤ 5 mm. The incidence of LLD perception is therefore higher than incidence of radiographically measured LLD, which is in agreement with previous studies [39]. Edeén et al. found the extent of LLD correlated with the awareness of LLD [40], but in the present study we found this to be true only if LLD was measured with the clinical relative method (Table 2). Results further indicate that overall awareness of LLD depends not on leg elongation per se (Table 3) but rather on the combined supratrochanteric and subtrochanteric causes of LLD, including preexisting contractures of lumbar spine, hip, and knee.

The importance of choosing the right outcome measure for a specific problem is also reflected in this study. While OHS clinical scores in our population did not show any significant correspondence with LLD or leg elongation values after adjustment for confounding factors, WOMAC scores in the identical group of patients produced different results. The question therefore arises whether a study analyzing LLD with a 12-item clinical score [22] can sufficiently capture this aspect of clinical problems. It should be noted that the study of Benedetti et al. [22] analyzed leg elongation (not the final LLD) as the primary studied variable and similarly to our results no significant correlation was found between leg elongation and subjective perception of discrepancy. Likewise, we found none of the radiographic parameters to be a significant predictor of overall patient satisfaction [23]. Although we observed significantly worse clinical outcomes with perceived LLD, the difference in mean clinical scores between patients with subjectively perceived LLD vs. unperceived LLD was approximately 10% (245 mm on the 0–2400 mm total WOMAC scale, 6 points on the 12–60 total OHS scale). In this regard we corroborate the conclusions of previous studies that the impact of postoperative LLD on the walking ability and overall clinical satisfaction is small [21–23].

The limitations of the present study include measurement errors of the methods used. It should be noted that in clinical methods (absolute or relative) tape measurements were performed with the nearest reading of 5 mm. Previous analyses have shown such measurement methods have sufficient reliability and accuracy with intraclass correlation coefficients >0.92 for interobserver reliability and >0.80 in comparison with CT scanogram [41]. Measurement errors of radiographic methods are in the range of 1–3 mm [25] and are thus considerably lower from the LLD and elongation amplitudes of the studied population. Another limitation of our study is the lack of a reference radiographic LLD measurement method that would take into account the entire leg-length of a subject. There are only a few studies published where full-leg radiographs [31] or CT scanograms [41,42] were used as the reference method for comparison with pelvic radiographic LLD measurements. We did not use full-leg radiographs or CT scanograms because they were not taken in the routine clinical setting and for ethical reasons we did not additionally expose subjects to unnecessary radiation. We did, however, use the interseteardrop line as the reference line for pelvic horizontality as it has been shown to have excellent agreement with leg-length measurements on full-
length radiographs [31]. The limitation of retrospective approach in the presented study is also small number of patients with preoperative clinical scores in contrast with high postoperative response rate for participation. Results regarding improvement of WOMAC score therefore have lower statistical power in comparison with postoperative WOMAC and OHS. Nevertheless, several previous studies have pointed out to the ‘ceiling’ effect of clinical scores in evaluation of THA patients and concluded that clinical scores more reliably describe postoperative state than clinical improvement [43,44].

To the knowledge of the authors this is the first study of LLD measurement in THA population to compare four different measurement methods and analyze their impact on different clinical outcome measures. We conclude the relationship between LLD and clinical outcome after THA depends considerably on the measurement method used, the parameter studied (leg elongation or final LLD) and the clinical outcome measure chosen. All these data should be taken into account when comparisons between different series of patients are made from the published data in the literature. The findings of this study have important implications for THA planning in clinical practice. They demonstrate that clinical LLD measurement methods are an indispensable factor in preoperative/postoperative patient assessment and thus THA planning cannot be based on pelvic radiographs alone. Leg elongation should be guided by clinical absolute LLD measurements while the patients’ postoperative self-perceived leg length discrepancy is best predicted by clinical relative LLD measurements. Therefore, the planned magnitude of intraoperative lengthening should be primarily based on preoperative absolute LLD and the final LLD testing after trial reduction of THA should be done with the relative LLD measurement method in order to take into account the intraoperative release of hip contracture. Within the <10 mm range of mean postoperative LLD in the studied series, the impact of subjectively perceived LLD on the overall clinical satisfaction is detectable but not considerable.

References