Heterotopic Ossification Following Arthroplasty for Femoral Neck Fracture

Comeau-Gauthier, Marianne MD, MSc; Zura, Robert D. MD; Bzovsky, Sofia MSc; Schemitsch, Emil H. MD, FRCSC; Axelrod, Daniel MD; Avram, Victoria MD, FRCSC; Manjoo, Ajay MD; Poolman, Rudolf W. MD, PhD; Frihagen, Frede MD, PhD, FRCSC; Heels-Ansdell, Diane MSc; Bhandari, Mohit MD, PhD, FRCSC; Sprague, Sheila PhD; the HEALTH Investigators


Abstract

Background:

Heterotopic ossification (HO) is a frequent complication following hip surgery. Using data from the Hip Fracture Evaluation with Alternatives of Total Hip Arthroplasty versus Hemiarthroplasty (HEALTH) trial, we aimed to (1) determine the prevalence of HO following total hip arthroplasty (THA) for femoral neck fracture in patients ≥50 years of age, (2) identify whether HO is associated with an increased risk of revision surgery within 24 months after the fracture, and (3) determine the impact of HO on functional outcomes.

Methods:

We performed a multivariable Cox regression analysis using revision surgery as the dependent variable and HO as the independent variable. We compared Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores between participants with and those without HO at 24 months.

Results:

Of 1,441 participants in the study, 287 (19.9%) developed HO within 24 months. HO was not associated with subsequent revision surgery. Grade-III HO was associated with statistically significant and clinically relevant deterioration in the total WOMAC score, which was mainly related to the function component of the score, compared with grade I or II.

Conclusions:

The impact of grade-III HO on the functional outcomes and quality of life after THA for hip fracture is clinically important, and HO prophylaxis for selected high-risk patients may be appropriate.

Level of Evidence:
Prognostic Level III. See Instructions for Authors for a complete description of levels of evidence.

Heterotopic ossification (HO) is the most frequent complication following primary total hip arthroplasty (THA). The reported frequency ranges from 5% to 87%, with an overall estimated prevalence of 30% to 40%\(^\text{1-4}\). It has been previously proposed that only severe HO alters the clinical course\(^\text{5-7}\). While severe HO is inarguably associated with a decreased range of motion\(^\text{1}\), disagreements persist regarding how this translates into functional outcomes\(^\text{8-10}\). Traditional risk factors and clinical outcomes following the development of HO have been determined mostly from studies of primary arthroplasty procedures for arthritis. Whether HO alters the outcomes of THA or hemiarthroplasty (HA) for femoral neck fracture to the same extent is uncertain.

Many authors have validated the use of nonsteroidal anti-inflammatory drugs (NSAIDs) for prophylaxis in populations undergoing elective hip surgery\(^\text{11-13}\). Recently, the use of aspirin as prophylaxis against venous thromboembolism has been suggested as an effective preventive measure against HO\(^\text{14-16}\), which would eliminate the need for other NSAIDs. The hip fracture population is relatively frail, older, and less fit to undergo unexpected surgery. The liberal use of NSAIDs in this population is not benign, and functional benefits from HO prophylaxis have not been clearly defined; therefore, appropriate patient selection for prevention is warranted.

Using data collected as part of the Hip Fracture Evaluation with Alternatives of Total Hip Arthroplasty versus Hemiarthroplasty (HEALTH) trial (ClinicalTrials.gov NCT00556842)\(^\text{17-18}\), we sought to better define HO following hip arthroplasty for hip fractures. HEALTH is a multicenter randomized controlled trial comparing THA and HA for the treatment of low-energy, isolated, displaced femoral neck fractures in patients \(\geq 50\) years old\(^\text{18}\). That trial was designed to evaluate the need for a secondary hip procedure within 24 months after the arthroplasty. In the present preplanned secondary analysis, we aimed to examine the prevalence of HO in this population, identify whether HO is associated with an increased risk of revision, and quantify the influence of HO on patients’ functional outcomes.

Materials and Methods

HEALTH Study Overview

Patients \(\geq 50\) years old presenting with an isolated, low-energy, displaced femoral neck fracture were enrolled at 80 participating sites in 10 countries and randomized to undergo THA or HA. Of the 1,495 patients in the HEALTH study, 1,441 (718 THAs and 723 HAs) were included in this secondary analysis. Participants were assessed clinically at 1 week, 10 weeks, and 6, 9, 12, 18, and 24 months postoperatively. The HEALTH trial was approved by the McMaster University Research Ethics Board (#06-151) and by the research ethics boards/institutional review boards of all participating clinical sites. The trial protocol and results have been published previously\(^\text{17-18}\).

HO Evaluation

After each participant had completed their 24-month follow-up, the central adjudication committee consisting of 3 orthopaedic surgeons, including a chair, reviewed all available
radiographs from scheduled and unscheduled clinical visits to identify the presence of HO. The chair independently reviewed all available radiographs for each participant who was identified as having HO and classified the severity of the HO according to the system developed by Brooker and colleagues (see Appendix A). We chose to use the Brooker classification (with the addition of grade 0 defined as the absence of HO) as it is the most widely used radiographic classification system for HO around the hip following THA or HA. All reviewers were blinded to clinical and functional outcomes. Given the poor interobserver and intraobserver reliability of the Brooker classification, Della Valle et al. suggested minimizing this source of discrepancy by pooling grades 0, I, and II to define low-severity HO and grades III and IV to define high-severity HO.

**Statistical Analysis**

The primary outcome measure in this analysis was the cumulative incidence of HO during the 24-month follow-up period. Additionally, we performed a multivariable Cox regression analysis to assess the impact of HO (the time-varying independent variable) on the risk of a revision procedure being performed within 24 months after fracture (the dependent variable). This analysis was adjusted for the following potential confounders: age, prefracture living status (institutionalized versus not institutionalized), prefracture ambulatory status (use of an assistive device to walk versus the ability to walk without an assistive device), American Society of Anesthesiologists (ASA) class (I/II versus III/IV/V), and type of arthroplasty (HA versus THA). After ensuring that the proportionality assumption was met, results were reported as hazard ratios (HRs) with 95% confidence intervals (CIs). All tests were 2-tailed with alpha = 0.05.

**Hip Function Assessment**

Using repeated-measures models for a 2-level independent analysis, we estimated the effect of HO on function, pain, and joint stiffness as measured with the self-administered Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) questionnaire (the dependent variable). With higher scores representing worse pain, stiffness, and function, the ranges for the WOMAC dimensions are 0 to 20 for pain, 0 to 8 for stiffness, and 0 to 68 for physical function. Patients were entered as random effects, while the presence and classification of HO at all visits (1 week, 10 weeks, and 6, 9, 12, 18, and 24 months) were included as time-varying independent variables in fixed effects. We included preinjury WOMAC scores as an adjustment variable. An autoregressive correlation structure was used to inform the model that each patient observation was expected to be correlated with its previous observation. Generalized linear models were created with Gaussian distributions. A threshold for the minimal clinically important difference (MCID) was set at 7 points based on previous literature. The results of the THAs and HAs were pooled as our original study showed that THA provided a clinically unimportant improvement in function and quality of life when compared with HA. Results were reported as adjusted mean differences (AMDs) with 99% CIs. All tests were 2-tailed with alpha = 0.01.

**Results**

*Cumulative Incidence of Heterotopic Ossification*
Of 1,441 patients (718 randomized to THA and 723 randomized to HA as part of the HEALTH trial) included in our analysis, 287 (19.9%) developed HO (Table I). Only 63 (4.4%) of the 1,441 patients developed high-grade (Brooker grade-III or IV) HO. The majority (218; 76.0%) of the 287 patients with HO had evidence of it on radiographs by 10 weeks postoperatively.

**HO and the Risk of Revision Surgery**

Of the 1,441 patients, 1,426 had sufficient data at 24 months for this analysis. Overall, there were 116 events of unplanned secondary procedures; however, none were performed for excision of HO. In the multivariable analysis, the diagnosis of HO was not associated with a higher risk of arthroplasty revision within 24 months after the femoral neck fracture (p = 0.70) (Table II).

**Hip Function Assessment**

Having low-grade HO (grade I or II) as compared with having no HO was not associated with a significantly higher mean total WOMAC score (Table III and Appendix B). Having high-grade HO (grade III or IV) as compared with no HO was associated with a significantly higher mean WOMAC total score (AMD = 8.23, 99% CI = 0.56 to 15.89; p = 0.006), indicating worse overall hip function (see Appendix C). Patients with high-grade HO (grade III or IV) as compared with the no-HO and low-grade-HO groups combined (grade 0, I, or II) had a significantly higher mean WOMAC total score (AMD = 8.66, 99% CI = 1.06 to 16.27; p = 0.003), WOMAC pain score (AMD = 1.54, 99% CI = 0.10 to 2.98; p = 0.006), and WOMAC functional score (AMD = 6.21, 99% CI = 0.26 to 12.16; p = 0.007), indicating worse function and pain at 24 months post-fracture (see Appendix D). However, only the AMD for the WOMAC total scores reached the MCID threshold of 7.

Grade-III HO was found to be associated with a clinically relevant and statistically significant deterioration in the WOMAC total score (AMD = 11.40, 99% CI = 2.18 to 20.61; p = 0.002) and function score (AMD = 8.41, 99% CI = 1.48 to 15.36; p = 0.002) as compared with low-severity HO (grade I or II) (Table IV and Appendix E). Following the same trend, high-grade HO (III or IV) was associated with a significantly worse WOMAC total score (AMD = 10.77, 99% CI = 1.70 to 16.63; p = 0.002) and function score (AMD = 7.95, 99% CI = 1.13 to 14.76; p = 0.003) compared with low-grade HO (I or II), with both score increases reaching the MCID (Table IV and Appendix F). Specifically, from the patient’s perspective, having low-severity HO (grade I or II) was not associated with a clinically important deterioration in functional outcome; however, grade-III HO was associated with a significant and clinically meaningful deterioration in function without affecting pain or joint stiffness.

**Discussion**

Using prospectively collected data from the HEALTH trial, we investigated the prevalence of HO following arthroplasty for hip fractures as well as the risk of revision surgery and the functional outcomes associated with this condition. Our findings confirm that the development of HO following THA or HA for a femoral neck fracture is fairly common, although the prevalence of 19.9% for any grade of HO at 24 months postoperatively in our study is lower than that in previous reports, in which the prevalence has usually ranged from 27.9% to 86.9%. The prevalence of severe HO (grade III or IV) was 4.4% (63 of 1,441) in
our study, in which only 5 patients (0.35%) had hip ankylosis (grade IV), which is similar to the previously reported prevalence of 3.6% to 7% for grades-III and IV HO following THA or HA. Considering the broad demographic features and the prospective nature of our study, we believe that it is likely representative of the prevalence of this complication in patients ≥50 years old in whom an isolated femoral neck fracture is treated by arthroplasty.

Revision surgery for HO is exceedingly rare, with reported rates of 0% to 0.40% and most of the procedures done for severely ankylosed hips (Brooker grade IV). In a consecutive series of 1,482 THAs, Alijanipour et al. reported no revision surgery for excision of HO. Similarly, Gofton et al. noted only 2 revision procedures for HO in their 10-year retrospective review of 1,087 primary THAs, in which the 2 patients were the only ones with grade-IV HO. This is consistent with our 0% revision rate for excision of HO and suggests that, even though we found an association between HO and hip-related reduction in functional outcomes, patients with HO following arthroplasty for a hip fracture are not necessarily more likely to require revision surgery.

Currently, it is generally accepted that low-grade HO (Brooker grades I and II) is not clinically relevant, which is consistent with our results. In contrast, there are considerable conflicting data regarding the clinical relevance of Brooker grade-III or IV HO. Several authors have found reduced walking ability, function, and range of motion (especially flexion) in patients with grade-III or IV HO, without an effect on pain scores. A meta-analysis by Neal showed Brooker grade-III and IV HO to have a negative correlation with range of motion and functional outcomes, whereas some authors concluded that only very severe HO (Brooker grade IV) is clinically relevant. In a study of 507 consecutive patients, Morrey et al. found a decreased arc of motion only in those affected by grade-IV HO and no statistically significant difference in functional outcomes or pain scores among all grades. Similarly, a prospective randomized study demonstrated no significant difference in outcomes among all of the Brooker grades with the exception of grade IV. Although authors of previous reports have hypothesized that only grade IV is clinically relevant, the results of our study have shown that patients with grade-III HO report a clinically meaningful deterioration in overall function compared with those with lower-severity HO (grades I and II), and this is not explained by increased pain or stiffness.

The results of this study have to be interpreted with caution since the sample-size calculation was based on a comparison of THA and HA with a primary outcome of unplanned secondary procedures, not on secondary outcomes such as HO. Also, grade-IV HO is exceedingly rare and was widely underrepresented in our study; thus, the comparison of the pooled results associated with grades III and IV may reflect functional outcomes and quality of life mostly related to grade III with very little contribution from grade IV.

To the best of our knowledge, this is the largest study reporting the prevalence of HO after arthroplasty in the setting of hip fractures. Fortunately, the prevalence of hip ankylosis (HO grade IV) remains rare. Additionally, our results demonstrate that grade-III HO is associated with worse functional outcomes; therefore, prophylaxis with acetylsalicylic acid or other NSAIDs may be recommended for carefully selected high-risk patients.

REFERENCES


TABLE I - Timing of Heterotopic Ossification Diagnosis After Fracture Surgery by Brooker Classification

<table>
<thead>
<tr>
<th>Brooker Classification (no. [%])</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 week</td>
<td>0 (0.0)</td>
<td>1 (0.3)</td>
<td>1 (0.3)</td>
<td>0 (0.0)</td>
<td>2 (0.7)</td>
</tr>
<tr>
<td>10 weeks</td>
<td>99 (34.5)</td>
<td>99 (34.5)</td>
<td>20 (7.0)</td>
<td>0 (0.0)</td>
<td>218 (76.0)</td>
</tr>
<tr>
<td>6 months</td>
<td>93 (32.4)</td>
<td>110 (38.3)</td>
<td>22 (7.7)</td>
<td>0 (0.0)</td>
<td>225 (78.4)</td>
</tr>
<tr>
<td>9 months</td>
<td>90 (31.4)</td>
<td>112 (39.0)</td>
<td>25 (8.7)</td>
<td>0 (0.0)</td>
<td>227 (79.1)</td>
</tr>
<tr>
<td>12 months</td>
<td>72 (25.1)</td>
<td>144 (50.2)</td>
<td>41 (14.3)</td>
<td>3 (1.0)</td>
<td>260 (90.6)</td>
</tr>
<tr>
<td>18 months</td>
<td>72 (25.1)</td>
<td>164 (57.1)</td>
<td>41 (14.3)</td>
<td>3 (1.0)</td>
<td>280 (97.6)</td>
</tr>
<tr>
<td>24 months</td>
<td>55 (19.2)</td>
<td>169 (58.9)</td>
<td>58 (20.2)</td>
<td>5 (1.7)</td>
<td>287 (100.0)</td>
</tr>
</tbody>
</table>

TABLE II - Association Between Heterotopic Ossification and Revision Surgery 24 Months Post-Fracture (N = 1,426; 116 Events)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>HR (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterotopic ossification: yes versus no</td>
<td>0.92 (0.58, 1.44)</td>
<td>0.70</td>
</tr>
<tr>
<td>Age</td>
<td>0.99 (0.97, 1.02)</td>
<td>0.52</td>
</tr>
<tr>
<td>Prefracture living status: institutionalized versus not institutionalized</td>
<td>0.93 (0.34, 2.54)</td>
<td>0.88</td>
</tr>
<tr>
<td>Independent ambulation: yes versus no (any aid used)</td>
<td>1.29 (0.84, 1.98)</td>
<td>0.25</td>
</tr>
<tr>
<td>ASA class: III/IV/V versus I/II</td>
<td>0.93 (0.64, 1.36)</td>
<td>0.70</td>
</tr>
<tr>
<td>Implant received: total hip arthroplasty versus hemiarthroplasty</td>
<td>0.95 (0.66, 1.37)</td>
<td>0.78</td>
</tr>
</tbody>
</table>
TABLE III - WOMAC Scores at 24 Months According to Brooker Classification of Heterotopic Ossification

<table>
<thead>
<tr>
<th>WOMAC Score†</th>
<th>Adjusted Mean Difference in Score at 24 Months (99% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 0 vs. Class I/II (N = 1,154)</td>
</tr>
<tr>
<td>Total</td>
<td>−1.54 (−5.73 to 2.65) 8.23 (0.56 to 15.89)‡</td>
</tr>
<tr>
<td>Pain</td>
<td>−0.48 (−1.26 to 0.30) 1.46 (−0.06 to 2.97)</td>
</tr>
<tr>
<td>Stiffness</td>
<td>0.09 (−0.31 to 0.49) 0.51 (−0.22 to 1.24)</td>
</tr>
<tr>
<td>Function</td>
<td>−1.26 (−4.55 to 2.03) 5.81 (−0.16 to 11.78)</td>
</tr>
</tbody>
</table>

*The mean difference was adjusted for baseline scores and obtained from the multilevel model.

†The minimal clinically important difference (MCID) was set at 7 points for the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

‡P < 0.01.

TABLE IV - Comparison of WOMAC Scores at 24 Months Between Low and High-Severity Heterotopic Ossification According to Brooker Classification

<table>
<thead>
<tr>
<th>Adjusted Mean Difference in Score at 24 Months (99% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I/II vs. Class III (N = 282)</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td>Stiffness</td>
</tr>
<tr>
<td>Function</td>
</tr>
</tbody>
</table>

*The mean difference was adjusted for baseline scores and obtained from the multilevel model.

†The minimal clinically important difference (MCID) was set at 7 points for the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

‡P < 0.01.