BMJ Open  Hip fracture evaluation with alternatives of total hip arthroplasty versus hemiarthroplasty (HEALTH): protocol for a multicentre randomised trial

Mohit Bhandari, P J Devereaux, Thomas A Einhorn, Lehana Thabane, Emil H Schemitsch, Kenneth J Koval, Frede Frihagen, Rudolf W Poolman, Kevin Tetsworth, Ernesto Guerra-Farfán, Kim Madden, Sheila Sprague, Gordon Guyatt, HEALTH Investigators

ABSTRACT

Introduction: Hip fractures are a leading cause of mortality and disability worldwide, and the number of hip fractures is expected to rise to over 6 million per year by 2050. The optimal approach for the surgical management of displaced femoral neck fractures remains unknown. Current evidence suggests the use of arthroplasty; however, there is lack of evidence regarding whether patients with displaced femoral neck fractures experience better outcomes with total hip arthroplasty (THA) or hemiarthroplasty (HA). The HEALTH trial compares outcomes following THA versus HA in patients 50 years of age or older with displaced femoral neck fractures.

Methods and analysis: HEALTH is a multicentre, randomised controlled trial where 1434 patients, 50 years of age or older, with displaced femoral neck fractures from international sites are randomised to receive either THA or HA. Exclusion criteria include associated major injuries of the lower extremity, hip infection(s) and a history of frank dementia. The primary outcome is unplanned secondary procedures and centre number. Data analysts and the HEALTH Steering Committee are blinded to the surgical arms, leaving the assessment of outcomes, patient quality of life, mortality and hip-related complications—both within 2 years of the initial surgery. We are using minimisation to ensure balance between intervention groups for the following factors: age, prefracture living, prefracture functional status, American Society for Anesthesiologists (ASA) Class and centre number. Data analysts and the HEALTH Steering Committee are blinded to the surgical allocation throughout the trial. Outcome analysis will be performed using a χ² test (or Fisher’s exact test) and Cox proportional hazards modelling estimate. All results will be presented with 95% CIs.

Ethics and dissemination: The HEALTH trial has received local and McMaster University Research Ethics Board (REB) approval (REB#: 06-151).

Results: Outcomes from the primary manuscript will be disseminated through publications in academic journals and presentations at relevant orthopaedic conferences. We will communicate trial results to all participating sites. Participating sites will communicate results with patients who have indicated an interest in knowing the results.

Trial registration number: The HEALTH trial is registered with clinicaltrials.gov (NCT00556842).

INTRODUCTION

The number of hip fractures globally is expected to rise to over 6 million by the year 2050.1 2 Owing to projected ageing of the...
population, the number of hip fractures is likely to exceed 500,000 annually in the USA and 88,000 in Canada over the next 40 years. By the year 2040, the estimated annual healthcare costs associated with hip fractures will reach $9.8 billion in the USA and $650 million in Canada. Hip fractures account for more hospital days than any other musculoskeletal injury and represent more than two-thirds of all hospital days due to fractures. In the elderly, hip fractures are associated with a 30% mortality rate of 1 year postinjury, and temporary and often permanent impairment of independence and quality of life. Worldwide, 4.5 million people are disabled from hip fractures yearly; the number of persons living with disability is expected to increase to 21 million in the next 40 years. The disability adjusted life-years lost as a result of hip fractures ranks in the top 10 of all causes of global disability.

Inconclusive clinical evidence
Femoral neck fractures may be either non-displaced (ie, very little separation at the fracture site, about one-third of femoral neck fractures) or displaced (ie, greater separation at the fracture site, about two-thirds of femoral neck fractures). Orthopaedic surgeons often treat non-displaced fractures with internal fixation and displaced fractures with arthroplasty, which is also known as joint replacement. We conducted a comprehensive literature search to identify all studies investigating total hip arthroplasty (THA; involves replacing the femoral head and the acetabulum) compared to hemiarthroplasty (HA; involves replacing the femoral head only) for displaced femoral neck fractures. We identified a recent meta-analysis of 14 studies (N=1890 patients). This meta-analysis demonstrated a lower risk of reoperation after THA compared with HA (relative risk 0.57, 95% CI 0.34 to 0.96); however, this effect was mainly driven by randomised controlled trials (RCTs) without concealed treatment allocation. THA consistently resulted in better hip function (Harris Hip Score, weighted mean difference 5.4, 95% CI 2.7 to 8.2) after follow-up intervals of 12–48 months. We also identified two additional meta-analyses that provide comparisons of THA and HA. Liao et al pooled RCTs of THA (N=403 patients) versus HA (N=425 patients), and found reductions in the risk of secondary procedures and wound infection with THA. Furthermore, Liao et al reported that the mobility rate in THA was better than in HA. Burgers et al performed a meta-analysis of eight RCTs (N=986 patients) that suggested similar rates of secondary procedures, 1 year mortality, major complications and minor complications with THA and HA, but large benefits in patient function with THA. Despite the benefits that THA confers regarding functional outcomes, secondary procedures and infection rates, both meta-analyses reported higher hip dislocation rates with THA.

In a survey of 298 surgeons who were members of the Orthopaedic Trauma Association or affiliated with AO International, 94–96% of surgeons agreed that arthroplasty is the preferred treatment option in patients over 80 years of age with displaced femoral neck fractures. In addition, 73% of surgeons surveyed preferred HA to THA. In summary, current evidence suggests substantial benefits of THA over HA through decreased pain and improved function, but also increased rates of dislocation associated with THA, with similar rates of secondary procedures and mortality with the two procedures. Despite the apparent benefits of THA over HA, surgeons more often prefer HA to treat displaced femoral neck fractures. Previous trials comparing these two approaches were limited by methodological issues; therefore, the current trial aims to determine the effectiveness of THA compared to HA using a large sample size and high-quality methodology.

Study objectives
Our primary objective is to assess the impact of THA versus HA on rates of unplanned secondary procedures within 2 years for individuals with displaced femoral neck fractures. Secondary objectives are:

1. To examine the effect of THA versus HA on health-related quality of life (Short Form-12, SF-12), functional outcomes and mobility (Western Ontario McMaster Osteoarthritis Index, WOMAC, and Timed Up and Go Test, TUG), and health outcome measures (EuroQol-5 Dimensions, EQ-5D).
2. To evaluate the effect of THA versus HA on mortality.
3. To evaluate the effect of THA versus HA on hip-related complications, including periprosthetic fracture, hip instability or dislocation, implant failure (loosening/subsidence and breakage), wound healing problems (including superficial/deep infection, wound necrosis), soft tissue problems (eg, pseudotumour), heterotopic ossification, abductor failure, implant wear and corrosion, osteolysis, neurovascular injury, decreased function and pain.

METHODS AND ANALYSIS
Overview of study design
HEALTH is a multicentre, concealed RCT of 1434 elderly patients who have sustained a displaced femoral neck fracture. We are using minimisation, a form of randomisation, to determine patient allocation. Surgeons across North America, Europe, Australia and Asia are participating. In conventional surgical hip fracture trials, all surgeons involved in the trial have performed both THA and HA based on the randomisation process. HEALTH utilises an expertise-based randomised trial design that allocates patients to surgeons with expertise in THA or HA (figure 1). Based on their expertise, surgeons perform either THA or HA. Surgeons who feel confident enough to perform both surgical strategies perform both THA and HA.
Patient selection

Eligibility criteria

The inclusion criteria are: (1) Adult men or women aged 50 years and older (with no upper age limit); (2) Fracture of the femoral neck confirmed with antero-posterior (AP) and lateral radiographs, or CT or MRI; (3) Displaced fracture that is not, in the judgment of the attending surgeon, optimally managed by reduction and internal fixation; (4) Operative treatment within 72 h of the patient being medically cleared for surgery; (5) Patient was ambulatory prior to fracture, though they may have used an aid such as a cane or a walker; (6) Anticipated medical optimisation for arthroplasty of the hip; (7) Provision of informed consent by patient or proxy; (8) Low energy fracture (defined as a fall from standing height); (9) No other major trauma (defined as an Injury Severity Score <17); and (10) Assurance that surgeons with expertise in both THA and HA are available to perform surgery.

The exclusion criteria are: (1) Patient not suitable for HA (eg, inflammatory arthritis, rheumatoid arthritis, pathological fracture (secondary to cancer) or severe osteoarthritis of the hip); (2) Associated major injuries of the lower extremity (eg, ipsilateral or contralateral fractures of the foot, ankle, tibia, fibula, knee or femur; dislocations of the ankle, knee or hip; or femoral head defects or fracture); (3) Retained hardware around the affected hip that will interfere with arthroplasty; (4) Infection around the hip (soft tissue or bone); (5) Patients with a disorder of bone metabolism other than osteoporosis (eg, Paget’s disease, renal osteodystrophy, osteomalacia); (6) Patients with a previous history of frank dementia that would interfere with assessment of the primary outcome (ie, secondary procedures at 2 years); (7) Likely problems, in the judgement of the investigators, with maintaining follow-up (eg, patients with no fixed address, report a plan to move out of town, alcohol abuse issues or intellectually challenged patients without adequate family support); and (8) Patients whose fracture occurred as a result of an act of violence.

Patient recruitment and screening

The first patient was randomised into the vanguard phase of this trial on 21 January 2009. After a brief pause in enrolment between the vanguard and definitive phase, the first definitive patient was randomised on 8 October 2013. Enrolment is ongoing at the time of publication and is expected to be completed by December 2016. All patients presenting to participating surgeons with a diagnosed femoral neck fracture amenable to arthroplasty are screened for participation in the HEALTH trial. Such patients are classified as: (1) Excluded (if they do not meet the eligibility criteria); (2) Missed (presumed eligible but missed due to error or staff availability); or (3) Included (eligible and randomised). Study personnel obtain informed consent from all eligible patients. If a patient lacks capacity and is deemed unable to consent, study personnel may obtain informed consent from the patient’s legally authorised representative and assent from the patient.

Randomisation

We are ensuring concealment of allocation by using a centralised 24 h computerised randomisation system that will allow internet-based randomisation. Patients are the unit of randomisation. To protect against prognostic imbalance between groups, we are using minimisation to ensure balance between intervention groups for several patient factors. The minimisation approach takes into account each preidentified prognostic variable and sums over the variables to allocate each patient to the treatment that will minimise the differences between groups for those prognostic variables. Unlike stratified
randomisation, minimisation will work towards minimising the total imbalance for all factors taken together instead of considering mutually exclusive subgroups (strata). Based on our international survey of surgeons and current evidence, we are minimising for the following prognostic factors: (1) Age (ie, 50–80 years or greater than 80 years); (2) prefracture living setting (ie, institutionalised or not institutionalised); (3) prefracture functional status (ie, using aid or independent ambulator); (4) American Society for Anesthesiologists (ASA) Class (ie, Class I/II or III/IV/V); and (5) participating centre.

Study interventions
Total hip arthroplasty
In the THA group, we are not including minimally invasive THA (ie, 2 incision approaches) or hinged prostheses or capture cups. To optimise feasibility and applicability of results, we did not standardise the surgical approach, including the use of cemented components, the implant manufacturer or femoral head size. We are documenting the manufacturer, implant material and bearing surface of the implant.

Hemiarthroplasty
Surgeons use modern implants for HA, excluding non-modular and non-canal filling unipolar implants, such as Moore’s and Thompson’s prostheses. We did not standardise the choice of modular unipolar versus bipolar HA nor whether implants are inserted with cement or with a press-fit design. We are documenting the manufacturer, implant material and bearing surface of the implant.

Standardisation of procedures and perioperative care
To ensure similar perioperative regimens, we recommend that participating centres standardise key aspects of preoperative and postoperative care. For preoperative care, we recommend standardisation of the following: (1) Preoperative antibiotic prophylaxis (eg, cephalosporins or equivalent coverage); (2) thromboprophylaxis (eg, oral anticoagulation drugs, thromboembolic deterrent stockings (TEDS), pneumatic compression boots or medical prophylaxis to be discontinued in sufficient time to allow for surgery as guided by international normalised ratio/partial thromboplasty time); and (3) medical consultation to optimise condition prior to surgery. For postoperative care, we recommend standardisation of the following: (1) Antibiotic prophylaxis (eg, cephalosporin or equivalent) for 24 h; (2) thromboprophylaxis with unfractionated heparin, low-molecular-weight heparin (LMWH), warfarin, antiplatelet agents or intermittent pneumatic compression boots; (3) weight-bearing as tolerated will be allowed as patients autoproptect the affected hip during rehabilitation. Post-surgery, patients are weight-bearing as tolerated and then advanced according to the attending surgeon’s best judgment; (4) 600 mg calcium by mouth daily and 1000 IU vitamin D per day (provided there are no contraindications), and further investigation and treatment of osteoporosis as recommended by a local osteoporosis expert/consultant; and (5) appropriate nutritional assessment with administration of oral micronutrient feeds as needed. Owing to a lack of evidence favouring a particular approach, we are recording but did not standardise the following: (1) Use of preoperative traction; (2) surgical delay; (3) type of anaesthetic (ie, general or regional); and (4) patient participation in physiotherapy and rehabilitation.

Study outcomes
Primary study outcome
The primary outcome is any unplanned secondary procedure within 2 years of the initial hip replacement surgery. A Central Adjudication Committee (CAC) will review each reported secondary procedure to determine that they are study events (ie, unplanned), and they will confirm the type of the procedure and the reason for the procedure (table 1).

Secondary study outcomes
Secondary study outcomes include: (1) Functional outcome and quality of life measured using self-administered and interview-administered questionnaires. (2) The effect of THA versus HA on mortality; (3) hip-related complications including periprosthetic fracture, hip instability or dislocation, implant failure (loosening/subsidence and breakage), wound healing problems (including superficial/deep infection, wound necrosis), soft tissue problems (eg, pseudotumour), heterotopic ossification, abductor failure, implant wear and corrosion, osteolysis, neurovascular injury, decreased function or pain.

Functional outcome and quality of life are measured using self-administered and interview-administered questionnaires. Functional outcome questionnaires include a generic health status measurement instrument (SF-12), a hip function and pain questionnaire (WOMAC), a health outcome measure (EQ-5D) and a functional mobility test (TUG).

The SF-12 questionnaire is a self-administered, 12-item questionnaire that measures health-related quality of life in eight domains. Both physical and mental summary scores can be obtained. Each domain is scored separately from 0 (lowest level) to 100 (highest level). The instrument has been extensively validated and has demonstrated good construct validity, high internal consistency and high test-retest reliability. It is frequently used in orthopaedics for evaluating fracture outcomes.

The WOMAC index is self-administered, 12-item questionnaire that measures health-related quality of life in knee and hip osteoarthritis using a battery of 24 questions. It is a valid, reliable and responsive measure of outcome and has been used in several studies involving a wide range of lower extremity conditions.

The EQ-5D is a comprehensive, compact health status classification and health state preference system. This questionnaire is widely used, and has demonstrated validity and sensitivity in many populations.
The EQ-5D is generalisable as it is widely used in Europe and will be useful for our definitive trial in which we will be including international sites.

The TUG involves observing the patient and documenting the time, in seconds, it takes for the patient to rise from a standard armchair, walk to a line on the floor 3 m away, turn, return and sits down again. The TUG has been used in many clinical contexts, including orthopaedics, rheumatology and for predicting geriatric falls.

**Classification of types and reasons for unplanned secondary procedures**

<table>
<thead>
<tr>
<th>Specific unplanned secondary procedures include</th>
<th>Classification of the reason for secondary procedures is as follows</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ Closed reduction of hip dislocation</td>
<td>▶ Treat a periprosthetic fracture</td>
</tr>
<tr>
<td>▶ Open reduction of hip dislocation</td>
<td>▶ Treat hip instability or dislocation</td>
</tr>
<tr>
<td>▶ Open reduction of fracture</td>
<td>▶ Treat infection—superficial</td>
</tr>
<tr>
<td>▶ Soft tissue procedure</td>
<td>▶ Treat infection—deep</td>
</tr>
<tr>
<td>▶ Insertion of antibiotic spacer</td>
<td>▶ Treat wound necrosis</td>
</tr>
<tr>
<td>▶ Full implant exchange</td>
<td>▶ Treat another wound healing problem</td>
</tr>
<tr>
<td>▶ Partial implant exchange—stem only</td>
<td>▶ Remove heterotopic ossification</td>
</tr>
<tr>
<td>▶ Partial implant exchange—head only</td>
<td>▶ Manage abductor failure</td>
</tr>
<tr>
<td>▶ Partial implant exchange—liner only</td>
<td>▶ Manage another soft tissue problem (ie, pseudotumour)</td>
</tr>
<tr>
<td>▶ Partial implant exchange—head and liner</td>
<td>▶ Correct implant failure—loosening or subsidence</td>
</tr>
<tr>
<td>▶ Partial implant exchange—acetabular component only</td>
<td>▶ Correct implant failure—breakage</td>
</tr>
<tr>
<td>▶ Partial implant exchange—acetabular component and head</td>
<td>▶ Treat implant wear</td>
</tr>
<tr>
<td>▶ Implant adjustment—re-orientation of the stem</td>
<td>▶ Treat osteolysis</td>
</tr>
<tr>
<td>▶ Implant adjustment—re-orientation of the acetabulum component</td>
<td>▶ Treat implant corrosion</td>
</tr>
<tr>
<td>▶ Implant removal with no replacement</td>
<td>▶ Improve function</td>
</tr>
<tr>
<td>▶ Excision heterotopic ossification</td>
<td>▶ Relieve pain</td>
</tr>
<tr>
<td>▶ Supplementary fixation</td>
<td></td>
</tr>
</tbody>
</table>

**Adjudication of study events**

The HEALTH CAC is comprised of five orthopaedic surgeons who specialise in hip surgery and have expertise in research methodology and experience with clinical trials. The CAC are reviewing: (1) Cases where fracture eligibility is in doubt; (2) postoperative X-rays to assess the technical placement of prostheses; (3) all reported secondary procedures and fracture-related complications to determine if the secondary procedure and/or fracture-related complication meets the requirements of the study criteria; and (4) cases of mortality to confirm the cause of death. All centres submit digital X-rays to the HEALTH Methods Centre. We post all relevant patient records devoid of personal identifiers (ie, case report forms and X-rays) on a specially designed, password-protected website for study adjudication. Adjudication occurs after patients have completed their 2-year follow-up. Any disagreements among the CAC members are resolved during conference calls. All decisions made by the Committee are final.

**Study follow-up**

All patients are followed for a period of 2 years. At each follow-up interval, patients’ health status and outcomes are recorded. In addition, during the 2-year follow-up visit, the surgeon documents any secondary procedure that may be planned for the patient. Figure 2 shows the schedule of events and assessments at each time point.

Our choice of a 2-year follow-up period is dictated by two factors. First, previous studies have reported that 75% of revision surgeries occur before 12 months for patients treated with arthroplasty for displaced femoral neck fractures. Thus, we can expect that the majority of revision surgeries will occur within 2 years. Since increasing follow-up to beyond 2 years will yield little additional information on secondary procedure rates, efficient use of resources dictates a 2-year follow-up. Second, a full 2 years of follow-up will provide sufficient time to assess any potential gains in function and quality of life provided by either surgical alternative.

**Protecting against sources of bias**

**Blinding**

While surgeons, patients and outcome assessors cannot be blinded to the surgical arms (ie, THA or HA), data analysts and the Steering Committee will remain blinded throughout the trial. Secondary procedures, the primary outcome, are objective and lack of blinding introduces minimal threats to validity. Additionally, the HEALTH trial design eliminates differential expertise bias by establishing a minimal threshold for experience, as well as ensuring that all surgeons, performing either THA or HA, are dedicated to and have sufficient expertise with the procedure.

**Surgeon expertise**

Surgeons participating in the HEALTH trial are asked to meet both of the following two criteria of expertise for either THA or HA: (1) Perform at least 50 procedures...
(either THA or HA) in their career (including residency experience in which they assumed responsibility for the procedure), and (2) perform at least five procedures per year. Surgeons who meet the threshold for both THA and HA will perform both procedures, if no overwhelming bias in favour of one procedure is evident. A surgeon is considered biased for an approach if they performed less than five cases of either procedure in their last 50 procedures for a displaced femoral neck fracture. Residents and fellows may perform the procedures under the supervision of a participating attending surgeon. The surgeon most responsible for the case must meet the threshold expertise criteria and must be present in the operating room for the critical aspects of the procedure (table 2). Our decision to set 50 procedures as the threshold for experience is based on previous studies examining learning curves and respective outcomes in patients undergoing THA.27–29 Our survey of surgeons, who are members of the Orthopaedic Trauma Association (OTA) or European AO International trauma centres, suggests that the threshold of HA is likely similar or slightly lower, given the perception of increased difficulty of THA.

Devereaux et al30 have outlined the advantages of this trial design which includes the following:

1. Elimination of differential expertise bias in which, in conventional designs, a larger proportion of surgeons are expert in one procedure under investigation over the other.

2. Differential performance, cointervention, data collection and outcome assessment are less likely than in conventional RCT.

With this approach, ethical concerns are reduced because all surgeries are conducted by surgeons with expertise and conviction in the concerned procedures.30

Maximising patient retention

Previous trials in hip fracture surgery have lost up to 50% of patients to follow-up.14 To avoid this problem, the strategies outlined in figure 3 are used to minimise the loss to follow-up. We have successfully used the majority of these strategies to maximise follow-up in other multicentre studies. Key features of this strategy include: (1) excluding individuals who are likely to present problems to the follow-up; (2) prior to hospital discharge, in addition to proving their own contact

<table>
<thead>
<tr>
<th>Radiographs &amp; Event Forms</th>
<th>Pre-Surgery</th>
<th>Surgery</th>
<th>Post Surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Screening</td>
<td>Enrolment</td>
<td>Baseline</td>
</tr>
<tr>
<td>Patient Contact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiographs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informed Consent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomisation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical Report</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Op Care</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinic or Telephone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-Up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary Procedure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse Event</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timed Up and Go</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WOMAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-5D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missed Follow-Up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early W/D</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 Schedule of events. *Complete forms when and/or if applicable.
information, each patient provides the name and address of alternate contacts who are likely to be aware of the patient’s whereabouts; (3) patients receive a reminder for their next follow-up visit from the clinical research coordinator; and (4) follow-up visits coincide with standard fracture clinic visits. Additionally, patients can complete some study visits over the phone provided no radiographs are required.

Minimising cointerventions and contamination
Any patients who cross over from one treatment group to the other will be analysed in the group to which they were allocated, maintaining the intention to treat approach of the analysis. Surgical cointerventions, such as general, neurosurgical or orthopaedic procedures, performed in addition to the arthroplasty procedure have the potential to confound outcomes. The standardisation of key aspects of preoperative and postoperative care, together with the expertise-based trial design, will minimise the use of cointerventions in study patients. Clinical sites record the details of any major additional procedures performed, as well as the use of medications that affect bone, such as bisphosphonates, vitamin D, calcium, hormone replacement therapy, selective oestrogen receptor modulators, calcitonin and anabolic steroid therapy.

Maximising protocol adherence
Given the inherent variability in practice patterns among orthopaedic surgeons, we believe it is important to ensure that surgeons adhere as closely as possible to the surgical management protocol. The feasibility of this large study involving multiple centres further depends on minimal changes from current practice. Our study protocol closely follows currently accepted practice in the management of patients with femoral neck fractures. We, therefore, anticipate high compliance with our protocol.

Statistical plan
Sample size determination
The choice of sample size is based on a comparison of THA versus HA for the primary outcome (unplanned secondary procedures). All statistical hypotheses will be two-sided. Levels of 0.05 for the primary and 0.01 for the secondary outcomes were chosen. Previous studies have reported secondary procedure rates in hip fracture patients treated with HA that have ranged from 4% to
Methods described by Collett. The goal is to calculate the proportional hazards model. The calculation is based on an approach to the primary analysis, which will use the Cox model.

Based on aggregate data from the pilot study, annual mortality rates of 15% and a loss to follow-up of 5% in each group were assumed, these rates applying to each of the 2 years of follow-up. Informed by the meta-analysis, the following assumptions were made: a 1-year risk of having a secondary procedure of 5% in the HA group and a corresponding risk of 1% in the second year. Various values of the relative risk reduction (HA vs THA) were then used to identify the specific value that would correspond to a cumulative risk difference between THA and HA of 5% after 2 years. This figure was identified in a survey of participating surgeons as the minimally important difference to clinicians (unpublished data). Annual event risks by group were converted into equivalent hazard rates, assuming for simplicity that the hazard rate would be approximately constant during each of the 2 years. It is estimated that approximately 72% of the group receiving THA and 76% of the group receiving HA will be event-free and available for further follow-up at the start of the second year. The sample size was increased to allow for a combined 7.6% crossover rate from the assigned to the alternate treatment, based on pilot data. These assumptions lead to a required total sample size of 1316 patients, which will yield an expected number of 96 secondary procedures. The associated relative risk reduction (RRR) is 0.45. These calculations were repeated after replacing the 5% event risk at 1 year for HA by 4% and 6% and leaving the other factors unchanged (Table 4). To account for potential surgeon level effects, the sample size has been further increased by 9% to 1434.

For the secondary outcomes, an important difference in SF-12 is considered to correspond to a moderate effect, as reported by Cohen, and to a minimally important difference in the SF-12 as reported by Ware. In both cases, the value is at least half the SD, equivalent to a 4-point difference in score. Specifying α level=0.05, β=0.20 (study power=80) and a SD of 8, a sample of at least 128 patients (64 per group) is required to detect the 0.1 point difference in the SF-12 as reported by Ware. In this study, at least 90 patients (45 per group) would be required to detect this difference (α level=0.05, β=0.20, σ=1). Previous studies have found a SD of 0.20 for the EQ-5D. To detect a difference of 0.10 (half the SD) with 80% power at α level of 0.05, this study requires a total of 128 patients (64 per group). Thus, in all circumstances, the desired sample size of 1434 patients will be sufficient to detect clinically meaningful differences in the secondary measures of outcome.

### Primary analyses
All outcome analyses will be performed by an intention-to-treat approach. To evaluate the effect of THA versus HA on time to unplanned secondary procedures (the event rate of having a secondary procedure of 5% in the HA group and a corresponding risk of 1% in the second year. Various values of the relative risk reduction (HA vs THA) were then used to identify the specific value that would correspond to a cumulative risk difference between THA and HA of 5% after 2 years. This figure was identified in a survey of participating surgeons as

### Table 3 Sample size calculations comparing total hip arthroplasty (THA) and hemiarthroplasty (HA)

<table>
<thead>
<tr>
<th>Study</th>
<th>THA reoperation rate</th>
<th>HA reoperation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>van den Berkem et al 81</td>
<td>5/110 (4.5%)</td>
<td>1/136 (0.7%)</td>
</tr>
<tr>
<td>Baker et al 82</td>
<td>4/36 (11.1%)</td>
<td>2/39 (5.1%)</td>
</tr>
<tr>
<td>Blomfeldt et al 83</td>
<td>4/56 (7.1%)</td>
<td>3/57 (5.3%)</td>
</tr>
<tr>
<td>Keating et al 84</td>
<td>6/63 (9.5%)</td>
<td>5/64 (7.8%)</td>
</tr>
<tr>
<td>Macaulay et al 85</td>
<td>1/16 (6.25%)</td>
<td>0/23 (0%)</td>
</tr>
</tbody>
</table>

Studies included in a meta-analysis comparing THA with HA for the outcome of reoperation at 1 year. THA versus HA combined effect: fixed and random, relative risk=1.67 (95% CI 0.86 to 3.24), p=0.13. Pooled event rates: THA fixed and random, incidence rate=7.8% (95% CI 5.2 to 11.6).

HA fixed, incidence rate=5.3% (95% CI 3.2 to 8.9).

HA random, incidence rate=4.9% (95% CI 2.6 to 9.2).

10% within 1 year, with a weighted pooled risk of 5.3% (95% CI 3.2% to 8.9%) in a fixed effect meta-analysis or 4.9% (95% CI 2.6% to 9.2%) using random effects. A pooled estimate from five randomised trials comparing THA with HA gave a relative risk of 1.67 (95% CI 0.86 to 3.24), p=0.13. Pooled event rates: THA fixed and random, incidence rate=7.8% (95% CI 5.2 to 11.6).

HA fixed, incidence rate=5.3% (95% CI 3.2 to 8.9).

HA random, incidence rate=4.9% (95% CI 2.6 to 9.2).

### Table 4 Sample size calculations comparing total hip arthroplasty (THA) and hemiarthroplasty (HA)

<table>
<thead>
<tr>
<th>HA 1 year event risk</th>
<th>Total patients required</th>
<th>Expected events (n)</th>
<th>RRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>4%</td>
<td>1123</td>
<td>72</td>
<td>0.50</td>
</tr>
<tr>
<td>5%</td>
<td>1316</td>
<td>96</td>
<td>0.45</td>
</tr>
<tr>
<td>6%</td>
<td>1435</td>
<td>108</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Sample size requirement for HEALTH Pivotal Trial (non-inferiority design). RRR, relative risk reduction.
primary outcome), a Cox proportional hazards model will be used with the following covariates: (1) Age (ie, 50–80 years or >80 years); (2) prefracture living setting (ie, institutionalised or not institutionalised); (3) prefracture functional status (ie, using aid or independent ambulator); (4) American Society for Anesthesiologists (ASA) Class (ie, Class I/II or III/IV/V); and (5) participating centre. Results will be reported as HRs with 95% CIs. Kaplan-Meier curves will be constructed.

Secondary analyses
In the secondary analysis, we will adjust for additional baseline factors when examining the effect of THA versus HA on time to unplanned secondary procedures (the primary outcome). This will be a Cox proportional hazards model which will include the same variables that are included in our primary analysis and additionally adjust for (1) gender, (2) surgical approach and (3) head size, by including these as independent variables. Results will be reported as HRs with 95% CIs.

A generalised linear model will estimate the effect of THA versus HA on quality of life (SF-12), function (WOMAC), health outcome (EQ-5D), and mobility (TUG) at follow-up using the following covariates that are included in our minimisation procedure: (1) Age (50–80 years or greater than 80 years); (2) prefracture living setting (ie, institutionalised or not institutionalised); (3) prefracture functional status (ie, using aid or independent ambulator); (4) ASA Class (ie, Class I/II or III/IV/V); and (5) participating centre.

Cox proportional hazards modelling will estimate the relative effect of THA versus HA on time to mortality and hip-related complications. Results will be reported as HRs with 95% CIs. Kaplan-Meier curves will be constructed.

Interim analyses
The approach to interim analyses is guided by a desire to avoid spuriously inflated estimates of treatment effect. A single interim analysis will be performed when 60% of the planned patient-years of follow-up have been accrued. The data analyst will present the results of these analyses to our independent Data Monitoring Committee. The committee will be guided by the O’Brien-Fleming stopping rule based on the primary outcome, which will maintain the overall specified type I error rate at 5% for the combined interim and final analyses. According to this rule, the required p value to declare a significant result in the interim analysis is 0.007 and at the final analysis, the required p value is 0.047. The rule is conservative, making it difficult to stop the trial early unless a large treatment effect is observed. We will only apply our stopping rule to the primary outcome. The secondary functional outcomes may demonstrate significance quickly due to the nature of the instruments, but we will not stop the study for that reason. No one other than committee members will be aware of the data on which the committee makes its decision and no one involved in the study will be aware of the content of their deliberations.

Data management
The case report forms (CRFs) are the primary data collection tool for the study. An Electronic Data Capture system (iDataFax) is being used to submit data to the Methods Centre located at McMaster University. On receipt of the data, the personnel at the Methods Centre make a visual check of the data and they query all missing data, implausible data and inconsistencies.

Sensitivity analyses
We will perform some sensitivity analyses to assess the robustness of the primary results. First, we will perform a competing risk analysis using Fine and Gray method to account for death as a competing risk. Second, we will use the random-effects analysis with a centre as a random-effect to account for the possibility of the centre effect or clustering within a centre. Lastly, based on the pilot results, we anticipate some crossovers during the trial—we will monitor these and perform some sensitivity analysis to assess their impact. All analyses will be performed using SAS V9.2 (Cary, North Carolina, USA).

Missing data
Patients who are lost or withdrawn from the study prior to the full 24 month follow-up will be censored at their last visit where evaluation of the primary end point was completed.

Data monitoring committee
The purpose of the Data Monitoring Committee (DMC) is to advise the HEALTH Investigators regarding the continuing safety of study participants. The DMC consists of the Chair, who is an orthopaedic surgeon, one biostatistician and two additional orthopaedic surgeons. All members are independent of the trial investigators, and have neither financial nor scientific conflicts of interest with the trial. Further details regarding the DMC can be found in the DMC Charter, available from the corresponding author.

ETHICAL CONSIDERATIONS
All patients (or their legally authorised representative) included in this study will sign a consent form that describes this study and provides sufficient information for patients to make an informed decision about their participation. All participating centres will submit the consent form with the protocol for review and approval by their local research ethics board (REB), or institutional review board (IRB) for the study. Centres will obtain written consent from every patient, using the REB/IRB-approved consent form, before that patient undergoes any study procedure. Any amendments to the study protocol which may affect the conduct of the study, or the potential safety of or benefits to patients...
will require a formal amendment to the protocol requiring approval by McMaster University’s REB and local research ethics boards for clinical sites. Any protocol amendments will be communicated to investigators, REB/IRB, trial participants and trial registries, as deemed necessary.

Information about study patients will be kept confidential and will be managed in accordance with the following rules: (1) All study-related information will be stored securely at the clinical site; (2) all study patient information will be stored in locked file cabinets and be accessible only to study personnel; (3) all CRFs will be identified only by a coded patient number and initials; (4) all records that contain patient names or other identifying information will be stored separately from the study records and can be identified only by the coded patient number and initials; and (5) all local databases will be password protected.

DISSEMINATION

Results from the primary manuscript will be submitted for publication regardless of whether or not there are significant findings. Every attempt will be made to ensure that the amount of time between completion of data collection and release of study findings are minimised.

Only the Methods Centre will have access to the full trial data set. Data for the primary publication will be analysed exclusively by the Methods Centre. Requests for access to the full trial data set for secondary publications are encouraged and can be initiated through a written request to Methods Centre personnel.

DISCUSSION

Previous orthopaedic trials have addressed the effect of THA versus HA on patient outcomes; however, the trials conducted to date are limited by small sample sizes, lack of concealed randomisation, differential expertise biases and most were single centre initiatives that lacked sufficient power to inform surgical practice. Orthopaedic surgeons appear to currently favour HA, although current evidence suggests superior patient-important outcomes with THA. The HEALTH trial will aim to resolve these controversies by establishing the effectiveness of each method of arthroplasty. This will have important clinical implications as each treatment is easily applicable and already in use in orthopaedic practice.

The HEALTH trial will address and overcome many of the limitations of previous orthopaedic hip fracture trials. First, the study sample size of 1434 patients is sufficient to detect small but important differences in outcomes and will ensure that our study objectives are met. Second, using an expertise-based RCT has many advantages over traditional orthopaedic hip studies, including eliminating differential performance and differential outcomes assessment, and reducing the number of crossovers and ethical concerns associated with randomisation. Furthermore, the HEALTH trial has stringent methodological safeguards against bias including: the use of a centralised system to randomise patients; blinding data analysts and the Steering Committee; standardisation and documentation of preoperative care, perioperative care and postoperative care; the use of strategies to limit loss to follow-up and adjudication of trial events by an independent CAC.

A key limitation of the HEALTH trial is that surgeons and patients cannot be blinded to the surgical arms. This leaves the assessment of outcomes and decisions to re-operate vulnerable to bias. We have limited this bias by using an objective primary outcome (unplanned secondary procedures) and by centrally adjudicating all primary outcome events using a group of independent orthopaedic surgeons. Additionally, we do not allow family members or friends to answer questionnaires on behalf of cognitively impaired patients; thus, some secondary outcomes will only apply to patients who are not cognitively impaired. Our trial is limited to shorter term outcomes—up to 2 years following the initial surgery. Therefore, our results may not apply when considering longer term outcomes.

The results of the HEALTH trial will be an important contribution to orthopaedic surgical literature and are likely to lead to changes in orthopaedic practice. Identifying the optimal approach to arthroplasty has the potential to improve the lives of hundreds of thousands of patients and to reduce the economic burden associated with hip fractures. If this trial shows that one approach is superior to the other, it will revolutionise the treatment of hip fractures globally and may potentially lead to the establishment of new clinical guidelines for treatment of displaced femoral neck fractures. In addition to the clinical impact, the HEALTH trial also has the potential to impact orthopaedic trial conduct. With this trial we will build collaborative relationships among countries and between clinical centres. This trial will also contribute to challenging the dogma that surgical trials are doomed to be confined to small single centre initiatives.

Author affiliations
1Department of Surgery, McMaster University, Hamilton, Ontario, Canada
2Department of Medicine, McMaster University, Hamilton, Ontario, Canada
3Boston University Medical Center, Boston, Massachusetts, USA
4Department of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, Ontario, Canada
5St Michael’s Hospital, Toronto, Ontario, Canada
6Department of Orthopaedic Surgery, Orlando Regional Medical Centre, Orlando, Florida, USA
7Orthopaedic Centre, Ulleval University Hospital, Oslo, Norway
8Department of Orthopaedic Surgery, Onze Lieve Vrouwe Gasthuis/University of Amsterdam, Amsterdam, the Netherlands
9University of Queensland, Queensland, Australia
10Department of Traumatology Orthopaedic Surgery and Emergency, University Hospital Vall d’Hebron, Barcelona, Spain
11Hamilton, Ontario, Canada
12McMaster University, Hamilton, Ontario, Canada


Canada Research Chair in Musculoskeletal Trauma which is unrelated to the present study (McMaster University, Hamilton, ON, Canada).

Competing interests None.

Ethics approval McMaster University Research Ethics Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.