Core Decompression with Bone Marrow Stem Cells for Osteonecrosis of Femoral Head

Fariba Farrokhi1, Sneha Priyadarshini Honnabovi1, Marisa Pavone1, Kamal AL-Eryani2, Oussama Abousamra3, Reyes Enciso4*

1Advanced graduate, Master of Science Program in Orofacial Pain and Oral Medicine, Herman Ostrow School of Dentistry, University of Southern California, Los Angeles, CA, USA
2Assistant Professor of Clinical Dentistry, Division of Periodontology, Diagnostic Sciences & Dental Hygiene, Herman Ostrow School of Dentistry, University of Southern California, Los Angeles, CA, USA
3Assistant Professor of Clinical Orthopaedic Surgery, Children's Hospital Los Angeles, University of Southern California, Los Angeles, CA, USA
4Associate Professor – Instructional, Division of Dental Public Health and Pediatric Dentistry, Herman Ostrow School of Dentistry, University of Southern California, Los Angeles, CA, USA

Article Info

Article Notes
Received: November 18, 2020
Accepted: December 22, 2020

*Correspondence:
*Dr. Reyes Enciso, Associate Professor - Instructional, Division of Dental Public Health and Pediatric Dentistry, Herman Ostrow School of Dentistry, University of Southern California, Los Angeles, CA, USA; Telephone No: (213) 821-6730; Fax No: (213) 740-8815; Email: renciso@usc.edu.

©2020 Enciso R. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License.

Keywords
Osteonecrosis
Femoral head
Core decompression
Bone marrow stem cells
Systematic review
Meta-analysis

Abstract

The authors conducted a systematic review and meta-analysis to determine if Core Decompression (CD) with Bone Marrow Stem Cells (BMSC) is more effective in treating Osteonecrosis of the Femoral Head (ONFH) compared to CD. Authors used Cochrane Library, EMBASE, PubMed, Web of Science, and hand-searched references through January 2020, identifying relevant randomized controlled trials (RCTs). Risk of bias was assessed with Cochrane's handbook. Fifty-four abstracts were screened, and eight RCTs (five at high and three at unclear risk of bias) with 432 patients were included. Meta-analyses found statistically significant improvement in Harris Hip Score (HHS) at 12 months (Difference in Means [DM]=10.065; 95% Confidence Interval [CI]=4.509 to 15.622; p<0.001) and pain intensity at 24 months (DM=-7.364; 95% CI=-12.113 to -2.615; p=0.002) in CD+BMSC group compared to CD alone although these results may not be clinically significant. Risk of Total Hip Replacement (THR) in patients receiving CD+BMSC was 33.4% lower than in CD group though not significant (RR=0.666; 95% CI=0.355 to 1.250; p=0.206). Though meta-analyses found the addition of BMSC to CD significantly improves clinical outcomes (HHS and pain intensity) compared to CD only, evidence was of moderate/low quality due to high risk of bias, imprecision, and small sample sizes. Further research is needed to confirm the results.

INTRODUCTION

Osteonecrosis of the femoral head (ONFH) is a term used to describe cellular death (necrosis) of the femoral bone tissue due to a lack of blood supply1. ONFH is a common condition, which usually starts unilaterally, but in more than half of the cases, it advances to the contralateral side within two years2. The blood supply to the femoral head can be interrupted by several different mechanisms, including trauma (femoral head and neck fractures, or surgical intervention to address specific hip pathologies), embolism, or thrombosis, as well as pressure or injury to the wall of the vessels, which could cause complete occlusion1,3. This interruption in the blood supply results in necrosis of the bone tissue, and collapse of the joint articular surfaces5. The pathological changes in ONFH include necrosis of hematopoietic cells and adipocytes, followed by necrosis of osteocytes1. These pathological changes in both the vasculature and bone cause bone remodeling, which may cause a subchondral fracture or bone collapse6.

Every year, in the United States, approximately 20,000 to 30,000
people develop ONFH. Men are more susceptible to this condition than women, and most affected individuals are between thirty to sixty years of age. Associated risk factors for ONFH include trauma, alcohol abuse, long-term use of corticosteroids, pregnancy, coagulopathy, organ transplant, inflammatory, and autoimmune disease. Up to 38% of osteonecrosis cases are due to oral or intravenous use of corticosteroids.

According to the staging system of ONFH by the Association Research Circulation Osseous (ARCO) originally presented in 1994 and most recently reviewed in 2019, there are four stages of osteonecrosis of the femoral head: stage I - X-ray is normal, but either magnetic resonance imaging or bone scan is positive; stage II—X-ray is abnormal but without any evidence of subchondral fracture, fracture in the necrotic portion, or flattening of the femoral head; stage III—a fracture in the subchondral or necrotic zone as seen on X-ray or computed tomography scans; stage IV—X-ray evidence of osteoarthritis with accompanying joint space narrowing, acetabular changes, and/or joint destruction. Patients in the early stages of ONFH (ARCO stages I-II) may not experience any hip pain. As the disease advances, patients develop a dull, throbbing pain in the buttocks or the groin area, and find it hard to stand up or put weight on the involved hip. The diagnosis of ONFH at an early stage is essential to prevent the disease’s progression to the later stages, which may result in bone collapse or subchondral fracture.

There are two types of treatments for ONFH, surgical and non-surgical. Non-surgical treatments include restricted weight-bearing, medications, and biophysical modalities. These treatments help patients in the early stages of the disease with improved hip function, pain relief, and possible prevention of radiographic subchondral fracture. Surgically, hip-preserving procedures are treatment modalities performed before any fracture or collapse of the femoral head, and they include core decompression (CD), non-vascularized, and vascularized bone grafts. CD is accomplished by drilling a large hole or several small holes in the femoral head; these holes will help relieve pressure and bring more vascularity to the area of the necrotic bone. CD is often combined with grafting procedures using different types of grafts, including autologous or allograft cancellous bone, free vascularized bone grafts with iliac or fibular bone, free cortical grafts, osteochondral grafts, and muscle-pedicle grafts. The grafts are often used to provide structural support for the bone remodeling process. Stem cell grafting has also been used to promote formation of new bone. These procedures could be beneficial for patients with the early stages of osteonecrosis before the disease process leads to collapse of the femoral head and eventually to total hip replacement (THR) or arthroplasty.

THR should be considered the last resort, especially in young and middle-aged adults with ONFH, because it usually does not last for the patient’s entire lifetime and needs future revisions.

Stem cells were found in the bone marrow and thus referred to as Bone Marrow Stem Cells (BMSCs). BMSCs include pluripotent hematopoietic stem cells, which give rise to platelets, white and red blood cells. These hematopoietic stem cells are valuable in treating different forms of blood cancers. Also, BMSCs are the source of Mesenchymal Stem Cells (MSCs), also known as mesenchymal stromal cells; these are multipotent stem cells that are able to give rise to bone and cartilage. It has been reported that Injections of BMSCs into the necrotic zone of the femoral head could improve the outcomes of the ONFH treatment.

In one of the earliest trials, bone marrow harvested from the iliac crest (autogenous graft) was concentrated through a centrifuge device and then used during core decompression. The stem cells in the bone marrow are capable of self-renewal and regeneration of different cells and can potentially repair damaged tissues by replacing the diseased cells with minimal risk of rejection or side effects. Since ONFH occurs due to vascular and bone pathology resulting in altered bone remodeling, the BMSCs can assist in repairing damaged tissues through angiogenesis and osteogenesis.

This systematic review aims to explore the efficacy of Bone Marrow Stem Cells with CD compared with CD alone in the femoral head.

METHODS AND MATERIALS

Research Question

This systematic review followed the Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement. The PICOS question was:

- **Study design:** Randomized Controlled Trials (RCTs).
- **Population:** Adult patients with ONFH.
- **Intervention:** Core decompression (with or without autologous bone graft) followed by the insertion of stem cells.
- **Comparison:** Core decompression (with or without autologous bone graft).
- **Outcomes:** Harris Hip Score (HHS), conversion to THR, hip survival rates, Visual Analog Scale (VAS), WOMAC (Western Ontario and McMaster Universities Arthritis Index) score, volume of necrotic lesion, collapse of femoral head, and clinical findings.
- **Setting:** Hospital or university clinical care centers.
Inclusion and exclusion criteria

Inclusions: Only RCTs on patients with diagnosed OFHN were included in this review. All patients received core decompression. The study population received bone marrow stem cells with or without bone graft. A comparison group received CD or CD and saline injection with or without bone graft.

Exclusions: Other study designs (open-label, not randomized, pilot, observational), case series and reports, editorials, opinions, literature reviews, systematic reviews, meta-analyses and clinical guidelines will be excluded. Articles published in a language other than English will also be excluded.

The electronic databases were searched using the strategies reported in Table 1.

Data extraction and management

MEDLINE via PubMed, Web of Science, the Cochrane Library and EMBASE databases were searched by the senior author (R.E.). Three review authors (M.P., F.F., S.P.H.) scanned the titles and abstracts of the articles retrieved from the application of the search strategy, and they acquired the full manuscript if the RCT met the inclusion criteria, or when a definite decision could not be made regarding inclusion or exclusion based on the title and abstract only. Disagreements over inclusion/exclusion were resolved by discussion with a fourth author (R.E.). If the study was rejected, the reason for exclusion was recorded. The full-text articles were analyzed for inclusion/exclusion, and relevant data were extracted independently by the same three review authors (MP, FF, SPH) using a previously prepared data extraction form. The three forms from the separate reviewers were then combined by the fourth author (R.E.). The form included for each study: the study design, funding, characteristics of research subjects (sample size, inclusion/exclusion criteria), interventions, control groups, and outcomes.

Assessment of risk of bias and quality of evidence

The assessment of the risk of bias in the included RCTs was consistent with the approach described in the Cochrane Handbook for Systematic Reviews of Interventions (risk of bias tool) and was undertaken independently and in duplicate (two out of three reviewers MP, FF, SPH reviewed each manuscript) as part of the data extraction process as described previously. Discrepancies were resolved by the senior author (R.E.). The quality of evidence assessment and the summary of the review findings were conducted with the software GRADEpro© (Grader©), in accordance with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group, and Cochrane handbook. Pooled results were considered a low sample size if the total number of subjects was <400.

Statistical Analyses

Only trials with similar interventions (CD+ stem cells vs. CD) and similar outcomes (i.e., post-treatment HHS, WOMAC score, VAS, or THR) were pooled into a meta-analysis. Treatment effects were expressed as differences in means (DM) in the change from the baseline in HHS, VAS, and WOMAC score, with 95% confidence intervals. Review authors calculated estimates of effect as risk ratios with 95% confidence intervals (CI) for the number of patients presenting with THR and the number of patients with the collapse of the femoral head. Subgroup analyses for 12 months and 24 months for HHS outcomes were conducted. Sensitivity analyses with only an overall low risk of bias will be conducted if appropriate. Statistical heterogeneity was tested with the Cochran Q test and the $I^2$ statistic.

Table 1: Electronic database search strategies up to 1/22/2020

<table>
<thead>
<tr>
<th>Electronic database</th>
<th>Search strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE via PubMed</td>
<td>(“Osteonecrosis”[Mesh] OR osteonecrosis OR (vascular necrosis) OR (bone infarction) OR (aseptic necrosis) OR (ischemic bone necrosis)) AND (“Hip”[Mesh] OR hip OR “femoral head” OR femur head) AND (“Stem Cell Transplantation”[Mesh] OR (stem cell) OR (bone marrow)) AND (decompression)</td>
</tr>
<tr>
<td>search strategy: (searched on 1/24/2019); re-run on 1/22/2020)</td>
<td>Language: limit to English</td>
</tr>
<tr>
<td>Species: limit to Humans</td>
<td>Article types: limit to Clinical Trials, Randomized Controlled Trials, Review, Systematic Reviews, Guideline, Meta-analysis, Practice Guideline</td>
</tr>
<tr>
<td>The Web of Science and The Cochrane Library search strategy: (searched on 1/24/2019); re-run on 1/22/2020</td>
<td>(osteonecrosis OR (vascular necrosis) OR (bone infarction) OR (aseptic necrosis) OR (ischemic bone necrosis)) AND (hip OR “femoral head” OR femur head) AND ((stem cell) OR (bone marrow)) AND (decompression)</td>
</tr>
<tr>
<td>EMBASE search strategy: (searched on 1/24/2019); re-run on 1/22/2020</td>
<td>#1. osteonecrosis OR (vascular necrosis) OR (bone infarction) OR (aseptic necrosis) OR (ischemic bone necrosis)</td>
</tr>
<tr>
<td>#2. hip OR “femoral head” OR (femur head)</td>
<td>#3. (stem cell) OR (bone marrow)</td>
</tr>
<tr>
<td>#4. Decompression</td>
<td>#5. #1 and #2 and #3 and #4</td>
</tr>
<tr>
<td>#6. #5 AND (“article’/it OR ‘article in press’/it OR ‘conference paper’/it OR ‘review’/it)</td>
<td></td>
</tr>
</tbody>
</table>

Page 32 of 47
Estimates of effect were combined with a random-effects model if there was heterogeneity (Q-test p-value<0.10) or with the fixed-effect model otherwise. Comprehensive Meta-analysis software version 3 (Biostat, Englewood, NJ, USA) was used to conduct statistical analyses.

RESULTS

Results of the search

Through the preliminary search strategy by database searching, we arrived at 606 references, and a further ten more records were discovered through other sources like searching references of included studies. After duplicate elimination, 515 references were studied individually by three review authors (MP, FF, SH), and based on the abstracts and titles, the references were narrowed to 54 pertinent manuscripts. The full texts of these 54 manuscripts were analyzed for inclusion individually by the same authors, and eight manuscripts were found relevant for inclusion. Exclusions were based on if they were not RCTs (n=10), not in English (n=3), studied different conditions (n=1), different interventions (n=6), or different outcomes (n=2), and if they were proceedings abstracts (n=2), opinions/CE articles (n=2), reviews (n=9), animal trials (n=2), protocols (n=1), case reports (n=1), and duplicates (n=7).

PRISMA flowchart shows a summary of our results (Figure 1). A re-run of the search conducted on 1/22/2020 did not produce any new RCTs needing inclusion in the review.

Study Characteristics

The features of the included RCTs are summarized in Table 2. In total, eight RCTs with 432 patients were included. In 2008 study by Hernigou et al., titled a prospective, randomized study, the hip with the smaller size ONFH was assigned to CD only, and the opposite side hip with the larger ONFH was allocated to CD and BMSC treatment creating a high risk of bias for randomization as it is not based on a randomized sequence generated by an outside statistician but based on the patient's disease characteristics. The word "randomization" is not mentioned.
in the trial by Rastogi et al.33, however, the authors explained that allocation concealment was performed with “sealed opaque envelopes which were opened on the operating table by the surgeon.” Making this trial at high risk of bias for randomization as it is not clearly stated how the authors conducted randomization. The corresponding author was contacted multiple times to get this information without response.

**Diagnosis:** Participants were diagnosed with ONFH stage via Association Research Circulation Osseous (ARCO) diagnostic criteria34, FICAT35, and Steinberg36. Included patients had a diagnosis of ARCO stage I-III16,28-33, Steinberg I-II32, and FICAT score I-III27 (Table 2). Three trials included only non-traumatic hips16,28,30 with two29,31, including hips with and without trauma (Table 2).

**Table 2: Eligible studies included in this systematic review**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Country</th>
<th>Total sample size (patients and hips)</th>
<th>Intervention group, sample size (n)</th>
<th>Disease Classification</th>
<th>Inclusion criteria</th>
<th>Gender</th>
<th>Age (mean ± SD, range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauzeur et al.16</td>
<td>Belgium</td>
<td>38 patients 46 hips</td>
<td>CD + BMSC: 23 hips CD + Saline: 23 hips</td>
<td>ARCO III</td>
<td>Aged 18 years or older; non-traumatic ONFH; With a surface collapse lower than 30% of the entire articular surface together with a dome depression of no more than 4 mm</td>
<td>27M/11F</td>
<td>48.0 ± 2.8</td>
</tr>
<tr>
<td>Hernigou et al.32</td>
<td>France</td>
<td>125 bilateral patients 250 hips</td>
<td>CD + BMSC: 125 hips</td>
<td>ARCO I-II Steinberg I-II</td>
<td>Bilateral, symptomatic hips</td>
<td>78M/47F</td>
<td>Mean age=36 Age range = [18-54]</td>
</tr>
<tr>
<td>Ma et al.27</td>
<td>China</td>
<td>39 patients 49 hips</td>
<td>CD + ABG+BMSC: 18 patients with 24 hips</td>
<td>Ficat I-III</td>
<td>Age between 18 to 55 years; Notable pain in the hip &amp; normal, minor, or mixed osteopenia or presence of crescent sign in plain radiograph</td>
<td>28M/11F</td>
<td>35.60 ± 8.05</td>
</tr>
<tr>
<td>Pepke et al.18</td>
<td>Germany</td>
<td>24 patients 25 hips</td>
<td>CD + BMSC: 11 Hips</td>
<td>ARCO II</td>
<td>Age over 18 years; no prior trauma</td>
<td>22M/3F</td>
<td>44.3±3.4</td>
</tr>
<tr>
<td>Rastogi et al.33</td>
<td>India</td>
<td>40 patients 60 hips</td>
<td>CD + BMSC: 30 hips</td>
<td>ARCO I-III</td>
<td>Age 18 and 50</td>
<td>19M/9F</td>
<td>34.67 ± 7.02</td>
</tr>
<tr>
<td>Sen et al.29</td>
<td>India</td>
<td>40 patients 51 hips</td>
<td>CD + BMSC: 26 hips</td>
<td>ARCO I-II</td>
<td>Moderate and extensive lesions of ARCO stage I and II</td>
<td>27M/13F</td>
<td>44.5±3.3</td>
</tr>
<tr>
<td>Tabata-bae et al.20</td>
<td>Iran</td>
<td>18 patients 28 hips</td>
<td>CD + BMSC: 14 hips in 14 patients</td>
<td>ARCO I-III</td>
<td>Patients with non-traumatic ONFH</td>
<td>19M/9F</td>
<td>31 ± 11.4</td>
</tr>
<tr>
<td>Zhao et al.14</td>
<td>China</td>
<td>100 patients 104 hips</td>
<td>CD + BMSC: 51 hips in 50 patients</td>
<td>ARCO Stage IIC to IIC</td>
<td>Age between 18 and 55 years; risk factors such as trauma, corticosteroid use, alcohol abuse, Caisson disease, and other mechanisms.</td>
<td>53M/47F</td>
<td>32.7±10.5</td>
</tr>
</tbody>
</table>

**Abbreviations:** CD: Core decompression; BMSC: Bone Marrow Stem Cells; ON: Osteonecrosis; ONFH: Osteonecrosis of the femoral head; ARCO: Association Research Circulation Osseous; UBM: Undifferentiated Bone Marrow; M: Male; F: Female; ABG: Autologous Bone Graft.
Bilateral trials: In Tabatabaee et al. 2005 30, ten patients had bilateral ONFH, and nine of those patients had one hip in the intervention and the other hip in the control group, whereas one patient with bilateral ONFH ended up with both hips in the control group. In Hernigou et al.32 all patients had bilateral ONFH, with both hips symptomatic. The magnitude of ONFH in both hips of each patient was evaluated with MRI, and the hip with the smaller size ON was assigned to core decompression, and the opposite side hip with the larger ONFH was assigned to core decompression with BMSC injection32.

Demographics: The minimum age reported in the studies was 18 years old, and the maximum age reported was 56 years old. All RCTs included male and female genders. RCTs were performed in Belgium, France, Germany, China, India, and Iran. All RCTs were published in English between 2012-2018. In five trials, baseline differences in demographics or ARCO stage between the intervention group and the control group revealed no statistical significance32,27,28,31,33.

Interventions

Table 3 describes in detail the interventions and comparison groups.

Intervention group: The interventions in the experimental groups included CD with the injection of bone marrow stem cells (BMSC)16,28-33, and one intervention group also had an autologous bone graft with a buffy coat27.

Control group: Control group underwent CD only28-32, CD with saline injection16, CD and autologous bone graft without a buffy coat27, or CD with unprocessed bone marrow injection, which was not run through the centrifuge and was not concentrated33.

Anesthesia: Half of the RCTs in this review used general anesthesia during the hip surgery16,28,30,32, three used epidural anesthesia27,29,33, and one did not indicate the type of anesthesia33.

BMSC: Bone marrow was harvested from various sites, including posterior iliac16,27,29, anterior iliac crest28,32, non-specific site of the iliac crest30,33, or subtrochanteric area in the femoral neck31. The volume of bone marrow harvested also varied in size from 10 mL21 to 400 mL16, with one RCT not stating the amount27. The amount of BMSC varied and was reported as an average of the concentration of mononuclear cells (10^6 mononuclear cells/mL31 = 3 × 10^9/mL27) or the volume of solution containing the mononuclear cells (10 mL28 to 50 mL16). The amount of time taken to concentrate the BMSCs was not stated in two trials16,28. In one study, the marrow was harvested under local anesthetic at one appointment; the cells were concentrated and injected during the hip surgery two weeks later31. The remainder of the trials harvested and concentrated the bone marrow in one visit for 6-30 minutes27,29,30,32,33.

Bone graft: Seven of the procedures used to insert the BMSCs were similar16,28-33; the concentrated bone marrow solution was injected into the CD site with a needle. Ma et al.27 harvested a cylinder of bone from the subtrochlear area for both treatment and control groups and coated it with a buffy coat, only the treatment group.

Closure: The material used to close the trochanter’s access site was reported as gel foam16,33, bone wax29,31, and bone plug30, while the remaining RCTs did not report the access’ closure27,28,32.

Post-op instructions: Only Hernigou et al.32 included clear postoperative non-weight bearing period, using crutches for 10 days then full weight-bearing. Two studies did not report post-op instructions16,28 and the remaining five reported no weight-bearing period ranging from 2 to 6 weeks followed by either partial weight-bearing for two weeks or full weight-bearing27,29,31,33. The length of the postoperative non-weight bearing period ranged from 10 days to 42 days (6 weeks).

Follow-up: Follow up periods ranged from 3 months to 25 years, with only one study not indicating the follow-up period32.

Outcomes

Harris Hip Score (HHS) is a clinician-based assessment that covers four domains: Pain (44 points), function (47 points), absence of deformity (4 points), and range of motion (5 points). The maximum score is 100: 90-100 is regarded an excellent result; 80-90 is good; 70-80 is fair, and any score below 70 is considered a poor result. HHS was reported in six trials28,29,31-33, however only three could be included in the meta-analysis28,29,33. Baseline and post-treatment averages and their standard deviations were reported in Sen et al.29, and averages and ranges (with standard deviations [SD] calculated as range/4) were reported in Hernigou et al.32. Only Pepke et al.28 showed a graph with mean ± Standard Error of the Mean [SEM]. Review authors used the graph to estimate means and SEM with a ruler; Standard deviation [SD] was calculated based on the standard formula, SD = SEM * sqrt(N), with N sample size and sqrt the root square. Two trials could not be used for meta-analysis due to lack of SD reported33 or because only percentage (%) increase from the preoperative score was reported and not means and SD of scores at baseline and post-treatment31.

THR Rate refers to the proportion of hips in each group that needed a “Total Hip Replacement,” with a higher rate indicating an unfavorable result. In six trials16,27,30,33 the number of hips that converted to THR were reported and in two16,29 the number of “Hip survivals” were reported. Hip survival is the number of hips that did not receive the THR and, therefore, the opposite of THR from which we calculated
the number of hips that converted to THR. A meta-analysis with eight included RCTs were conducted.\(^{16,27-33}\).

**VAS:** The visual analog scale (VAS) is a subjective measure for acute and chronic pain, in which scores are recorded by making a mark on a line that represents a scale between “no pain” and “worst pain.” VAS scores before and after intervention were reported in five trials.\(^{16,27,28,30,32}\).

For Pepke et al.\(^{28}\) which used a 0-10 VAS scale, four other trials used a VAS scale of 0-100.\(^{16,27,30,32}\) Review authors converted the numerical scores to a 0-100 scale by multiplying by 10 the 0-10 scale for comparison in Pepke et al.\(^{28}\). Higher the score on the VAS scale, worse is the pain reported.

**WOMAC:** The Western Ontario and McMaster Universities

---

### Table 3: Description of interventions, comparison groups

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention groups and sample size</th>
<th>Anesthetic</th>
<th>Bone marrow harvest site</th>
<th>Volume harvested bone marrow/Volume Concentrated BMSC</th>
<th>Bone marrow Concentration period</th>
<th>Type of stem cells</th>
<th>Bone graft and other implantation details</th>
<th>Control therapy and sample size</th>
<th>Control therapy details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauzeur et al.(^{16})</td>
<td>CD + BMSC: 23 hips</td>
<td>General anesthesia</td>
<td>posterior iliac crest</td>
<td>400 mL / 50 mL injected</td>
<td>Not stated</td>
<td>BMSC</td>
<td>Gel foam absorbable gelatin sponge sealed trephine access both groups</td>
<td>CD + Saline 23 hips</td>
<td>Sham skin incision posterior iliac crest and 50 mL of saline injected</td>
</tr>
<tr>
<td>Hernigou et al.(^{32})</td>
<td>CD + BMSC: 125 bilateral hips in 125 patients</td>
<td>General anesthesia</td>
<td>Anterior iliac crest</td>
<td>152 ± 16 mL / 20 mL injected</td>
<td>&lt; 30 min</td>
<td>BMSC</td>
<td>Intervention in hip with larger ONFH lesion</td>
<td>CD: 125 hips</td>
<td>CD of hip with smaller ONFH lesion</td>
</tr>
<tr>
<td>Ma et al.(^{27})</td>
<td>CD + BMSC + Bone cylinder harvested from subtrochanteric area 39 patients 49 hips</td>
<td>Continuous epidural anesthesia</td>
<td>Posterior iliac crest</td>
<td>Not stated / Buffy coat solution, from aspirated bone marrow 3 x 10^10 nucleated BM cells</td>
<td>at least 12 min</td>
<td>BMSC</td>
<td>Bone cylinder harvested from subtrochanteric area had buffy solution dripped onto bone surface and BMSC applied to bone cylinder over buffy coat</td>
<td>CD + Bone cylinder graft 21 patients with 25 hips</td>
<td>CD + Bone cylinder graft without buffy coat or BMSC applied to bone</td>
</tr>
<tr>
<td>Pepke et al.(^{28})</td>
<td>CD + BMSC: 11 Hips</td>
<td>General anesthesia</td>
<td>Ventral (Anterior) iliac crest</td>
<td>200-220 mL / 10 mL injected</td>
<td>Not stated</td>
<td>BMSC</td>
<td>CD: 14 Hips</td>
<td>CD only</td>
<td></td>
</tr>
<tr>
<td>Rastogi et al.(^{29})</td>
<td>CD + BMSC: 30 hips</td>
<td>Spinal epidural anesthesia</td>
<td>Iliac crest</td>
<td>60–70 mL/5 mL with 1.1 x 10^6 mononuclear cells</td>
<td>1 hour</td>
<td>BMSC</td>
<td>Trephine access plugged via gel foam</td>
<td>CD + UBM: 30 hips</td>
<td>30–50 mL Unprocessed bone marrow aspirate injected into CD site</td>
</tr>
<tr>
<td>Sen et al.(^{29})</td>
<td>CD + BMSC: 26 hips</td>
<td>Epidural anesthesia</td>
<td>Posterior iliac crest</td>
<td>120 - 180 mL/5 x 10^6 mononuclear cells keeping CD34+ cell count &gt; 5 x 10^7 count</td>
<td>2 hours</td>
<td>BMSC</td>
<td>Trephine access sealed with bone wax</td>
<td>CD: 25 hip</td>
<td>CD Only</td>
</tr>
<tr>
<td>Tabatabaee et al.(^{30})</td>
<td>CD + BMSC: 14 hips in 14 patients</td>
<td>General anesthesia</td>
<td>Iliac crest</td>
<td>200 mL/mean mononuclear cells count was at least 2 million cells/mL</td>
<td>5-10 min</td>
<td>BMSC</td>
<td>Trephine access sealed with allograft bone plug</td>
<td>CD: 14 hips in 13 patients</td>
<td>CD Only</td>
</tr>
<tr>
<td>Zhao et al.(^{31})</td>
<td>CD + BMSC: 51 hips in 50 patients</td>
<td>Not stated</td>
<td>Sub trochanteric area site</td>
<td>10 mL/2 mL solution containing 2x10^6 BMSC</td>
<td>2 weeks in vitro</td>
<td>BMSC</td>
<td>Trephine access sealed with bone wax</td>
<td>CD: n = 53 hips in 50 patients</td>
<td>CD Only</td>
</tr>
</tbody>
</table>

**Abbreviations:** CD: Core decompression; BMSC: Bone Marrow Stem Cells; UBM: Undifferentiated Bone Marrow.
Arthritis Index (WOMAC) used in the evaluation of hip and knee osteoarthritis is a self-administered questionnaire of 24 items divided into subscales of pain, stiffness and physical function. Higher scores on the WOMAC index signify worse pain, stiffness, and functional limitations. WOMAC sub-scales A and B (functional ability) were reported in Hauzeur et al.\textsuperscript{16} and three studies\textsuperscript{27,30,32} reported WOMAC average range. Two trials reported baseline and post-treatment WOMAC averages and SD at 24 months and were included in one meta-analysis\textsuperscript{27,30}. However, Hernigou et al.\textsuperscript{32} reported means and SD for those hips without collapse and without THA, hips with a primary THA and hips with THA revision separately at 25 years and this RCT was excluded from the meta-analysis.

The collapse of the femoral head and disease progression:
The surface collapse extension was measured with three different methods precluding a meta-analysis for this outcome. It was measured radiologically and expressed as a percentage of the whole articular surface with <15%, 15-30%, >30% involvement\textsuperscript{16}. Radiographic progression was interpreted as a progress in the Steinberg stage\textsuperscript{35}. In Ma et al.\textsuperscript{27}, the number of hips with progression from Ficat I to stage II and from stage II to III/IV were reported. A higher Steinberg or Ficat stage indicates a negative outcome.

Femoral head necrosis volume: The volume of the necrotic zone was evaluated with an MRI and reported as mm\textsuperscript{3}\textsuperscript{28}, modified Kerboul angle\textsuperscript{33} or volume of low signal intensity zone in the femoral head on the MRI\textsuperscript{31}. In the modified Kerboul angle’s method, the amount of necrotic lesion is measured on coronal and sagittal MRI T1-weighted images\textsuperscript{36}. The lesions were divided into small (Kerboul <160), medium (Kerboul 160–200), and large (Kerboul >200) based on the angle, and a hip was considered as having a favorable outcome if there was a reduction in the necrotic angle. The necrotic volume or low signal intensity zone in the MRI images acquired was measured using image analysis software. A meta-analysis could not be performed for femoral head necrosis volume due to the heterogeneity of the outcomes.

Lequesne index: The clinical symptoms of the joint were calculated by the Lequesne index only in Ma et al.\textsuperscript{27} and a meta-analysis could not be conducted. The Lequesne index of severity for osteoarthritis of the hip (LIJOH) is a subjective questionnaire\textsuperscript{37}. Patients’ responses in categories of pain or discomfort, maximum distance walked and activities of daily living are recorded on a scale from 0 to 24, and greater number indicates worse disability.

Adverse events:
Only Hauzeur et al.\textsuperscript{16} reported adverse effects which included pain at the great trochanter, pain at the iliac crest, fever for less than 24 hours with negative bacteriological investigations, and nausea. Another RCT\textsuperscript{27} stated that there were no surgical complications, such as bone marrow harvest site infection, surgical site infection, nor bleeding after the operation. Three trials reported no adverse reactions or complications\textsuperscript{28,30,31} and two\textsuperscript{29,32} did not mention any adverse events (Table 4).

### Table 4: Description of side effects

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention groups and sample size</th>
<th>Adverse events reported</th>
<th>I (## hips)</th>
<th>C (## hips)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauzeur et al.\textsuperscript{16}</td>
<td>CD + BMSC: 23 hips</td>
<td>Pain at the great trochanter, Pain at the iliac crest, Fever for less than 24 hours with negative bacteriological investigations Nausea</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Hernigou et al.\textsuperscript{32}</td>
<td>CD + BMSC: 125 bilateral patients</td>
<td>None stated</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Ma et al.\textsuperscript{27}</td>
<td>CD + ABG + BMSC: 18 patients with 24 hips</td>
<td>None reported</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pepke et al.\textsuperscript{28}</td>
<td>CD + BMSC: 11 hips</td>
<td>No side effects</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rastogi et al.\textsuperscript{31}</td>
<td>CD + BMSC: 30 hips</td>
<td>No intra-op or post-op complications</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sen et al.\textsuperscript{29}</td>
<td>CD + BMSC: 26 hips</td>
<td>Not reported</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tabatabaee et al.\textsuperscript{30}</td>
<td>CD + BMSC: 14 hips in 14 patients</td>
<td>No Adverse reaction reported</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zhao et al.\textsuperscript{31}</td>
<td>CD + BMSC: 51 hips in 50 patients</td>
<td>No complications were observed during the administration of anesthetics or after the operation in patients from the two treatment groups</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Abbreviations:** CD: Core decompression; BMSC: Bone Marrow Stem Cells.
Co-interventions

No co-interventions stated.

Risk of bias in included studies

The risk of bias graph (Figure 2) summarizes that four trials were considered unclear\textsuperscript{16,27,30,31} and four were high risk\textsuperscript{28,29,32,33} (Table 5).

Effects of interventions

Table 6 shows the outcomes reported in the included studies.

### Table 5: Summary of assessment of risk of bias

<table>
<thead>
<tr>
<th>Study</th>
<th>Random Seq. Generation</th>
<th>Allocation Concealment</th>
<th>Blinding</th>
<th>Incomplete Outcome Data</th>
<th>Selective Reporting</th>
<th>Other potential bias</th>
<th>Overall Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauzeur et al.\textsuperscript{16}</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>UNCLEAR</td>
</tr>
<tr>
<td>Hernigou et al.\textsuperscript{32}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>HIGH</td>
</tr>
<tr>
<td>Ma et al.\textsuperscript{27}</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>UNCLEAR</td>
</tr>
<tr>
<td>Pepke et al.\textsuperscript{28}</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HIGH</td>
</tr>
<tr>
<td>Rastogi et al.\textsuperscript{31}</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HIGH</td>
</tr>
<tr>
<td>Sen et al.\textsuperscript{29}</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>HIGH</td>
</tr>
<tr>
<td>Tabatabaee et al.\textsuperscript{30}</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>UNCLEAR</td>
</tr>
<tr>
<td>Zhao et al.\textsuperscript{31}</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>UNCLEAR</td>
</tr>
</tbody>
</table>

Legend: (+) High risk of bias; (-) Low risk of bias; (?) Unclear risk of bias

### Table 6: Results reported in the included studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Statistically significant differences between CD+BMSC and CD groups at follow-up</th>
<th>No statistically significant differences between groups (p&gt;0.05)</th>
<th>Conclusions (== No significant difference; &gt;&gt; better; &lt;&lt; worse)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauzeur et al.\textsuperscript{16}</td>
<td>N/A</td>
<td>No significant differences in:</td>
<td>CD+BMSC == CD+saline</td>
</tr>
<tr>
<td></td>
<td>• Significant decrease in the level of pain (p&lt;0.01) and joint symptoms (p&lt;0.03) observed in the BMSC group than CD group after 12 months.</td>
<td>• THR requirements,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 76% in CD group underwent THR compared to 24% in BMSC group (p &lt; 0.0001).</td>
<td>• clinical tests,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Hips undergoing only CD were approximately 3 times more likely to undergo a primary THR (odds ratio: 10; p &lt; 0.0001)</td>
<td>• radiological evolution at 24-month follow-up</td>
<td></td>
</tr>
<tr>
<td>Hernigou et al.\textsuperscript{32}</td>
<td></td>
<td></td>
<td>CD+BMSC &gt;&gt; CD</td>
</tr>
</tbody>
</table>
• Significant difference in pain relief and improved hip function following CD+BMSC after 3 months and 2 years. (VAS & Lequesne index) 33.3% of hips in control group deteriorated to the next stage after 24 months and 8% in the treatment group had further deterioration (p <0.05)  

CD+ABG+BMSC >> CD

Ma et al.27

No significant differences in:
• clinical outcomes (VAS and HHS)
• Radiographical outcome
• Interval from CD with or without BMSC application
• THR  

CD+BMSC == CD

Pepke et al.28

CD+BMSC >> CD+UBM

Rastogi et al.33

CD+BMSC >> CD+UBM

Sen et al.29

HHS:
• At 12 months: Statistically significant differences in HHS (p =0.016) and its domains (pain, function, deformity and motion) between groups.
• At 24 months: a statistically significant difference (p =0.05) in pain and deformity domains.

Clinical findings:
• Better clinical success noted in group CD+BMSC vs CD (p = 0.042).
• Hips with poor pre-operative HHS (< 70) and initial radiographic changes had significantly better hip survival CD+BMSC (p=0.0082) versus CD (p=0.0326)
• Hips with initial MRI edema had significantly better survival in CD+BMSC group than CD group at 12 months (p=0.019) and 24 months (p=0.047)
• Post-traumatic hips always had better outcome than non-traumatic hips  

No statistically significant difference in:
• HHS at 24 months (p = 0.09)

CD+BMSC == CD at 12 months (HHS) and in hips with poor pre-operative HHS <70, edema or post-traumatic hips

Tabata-baee et al.30

N/A

CD+BMSC >> CD

Zhao et al.31

CD+BMSC treatment significantly improved:
• HHS (p<0.05)
• decreased the volume of femoral head low signal intensity zone in the hips preoperatively classified at stage IC, IIB, and IIC (p<0.05, respectively; stage IIA, p=0.06)

No statistically significant difference in:
• HHS at 24 months (p = 0.09)

N/A

CD+BMSC >> CD

Abbreviations: CD: Core decompression; BMSC: Bone Marrow Stem Cells; ABG: Autologous Bone Graft; UBM: Undifferentiated Bone Marrow; THR: Total Hip Replacement; VAS: Visual Analogue Scale; HHS: Harris Hip Score; MRI: Magnetic Resonance Imaging; WOMAC: Western Ontario and McMaster Universities Arthritis Index.

Meta-analyses

HHS at 12 and 24 months

This review found a statistically significant favorable trials. In summary, six reported favorable results for CD + BMSC compared to control group27,29-33 and two found no significant differences16,28. Table 7 shows the THR rate in all included studies.
<table>
<thead>
<tr>
<th>Study</th>
<th>Time point</th>
<th>THR (%) Intervention Group</th>
<th>THR (%) CD Group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauzeur et al.</td>
<td>24 months</td>
<td>15/23 (65%)</td>
<td>15/23 (65%)</td>
<td>1.000</td>
</tr>
<tr>
<td>Hernigou et al.</td>
<td>25 years</td>
<td>30/125 (24%)</td>
<td>95/125 (76%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ma et al.</td>
<td>24 months</td>
<td>2/25 (8%)</td>
<td>8/24 (33%)</td>
<td>0.071</td>
</tr>
<tr>
<td>Pepke et al.</td>
<td>24 months</td>
<td>4/11 (36%)</td>
<td>6/14 (43%)</td>
<td>0.799</td>
</tr>
<tr>
<td>Rastogi et al.</td>
<td>24 months</td>
<td>0/30 (0%)</td>
<td>3/20 (10%)</td>
<td>0.198</td>
</tr>
<tr>
<td>Sen et al.</td>
<td>12 months</td>
<td>18/26 (69%)</td>
<td>14/25 (56%)</td>
<td>0.552</td>
</tr>
<tr>
<td>Tabatabaei et al.</td>
<td>24 months</td>
<td>0/14 (0%)</td>
<td>3/14 (21%)</td>
<td>0.198</td>
</tr>
<tr>
<td>Zhao et al.</td>
<td>60 months</td>
<td>0/50 (0%)</td>
<td>5/44 (11%)</td>
<td>0.087</td>
</tr>
</tbody>
</table>

**Summary of the evidence and quality of the findings (GRADE)**

Only trials reporting similar outcomes were pooled into meta-analyses. Sensitivity analyses with only overall low risk of bias RCTs were planned but not conducted due to the overall unclear/high risk of bias.

Due to unclear or high risk of bias, the quality of the evidence was moderate for the change in VAS pain in five trials with a total of more than 400 subjects in the meta-analysis. The quality of the evidence was low due to the risk of bias and imprecision (the small sample sizes in each meta-analysis < 400 participants) for THR rate with 7 RCTs, change in HHS at 12 months with 3 trials, and HHS at 24 months and change in average WOMAC score with 2 trials (Table 8).

**DISCUSSION**

**Summary of the main results in individual trials**

In this systematic review, six trials16,27-31,33 indicated use of bone marrow aspirate cells significantly improved the clinical outcomes in hips undergoing CD + BMSC compared to CD only. Two trials27,32 indicated the use of bone marrow aspirate cells did not improve the clinical outcome in the hips treated via CD + BMSC compared to CD only (Table 6).

**Summary of meta-analyses**

**Harris Hip Score**

Meta-analysis of HHS findings was based on three RCTs28,29,32 after 12 months, and two trials28,29 after 24 months. For both intervals, we found a statistically significant improvement in HHS in the intervention over the CD group by an average of 10 points higher (scale 0-100) at 12 months (p<0.001) and about 12 points higher at 24 months (p=0.026). Though this improvement appears to be statistically significant, it may not be clinically significant as it is less than 20 points. According to Singh et al.38, patients with ‘lack of improvement’ (defined as clinical improvement of HHS by 1-20 points from preoperative score to two years after THR), were associated with increased hazards of revision surgery compared to patients with an improvement of more than 50 points in HHS scores.
Figure 3: Meta-analyses. Change from baseline in HHS at 12 months (a) and 24 months (b); Total hip replacement rate (c); change in VAS pain from baseline at 24 months (d); change from baseline in average WOMAC score at 24 months (e).
Several systematic reviews in the past few years have compared the efficacy of BMSC with CD to that of CD alone. The most recent review in 2019 by Wang et al.39 came to the same conclusion as this review, and reported a significant difference in HSS scores at 12 months (WMD or weighted mean difference)= 4.80, 95% CI: [2.22, 7.39]; p=0.000), and at 24 months (WMD= 4.9, 95% CI: [3.06, 6.74]; p= 0.000), between the two groups. Our results are also in agreement with the systematic review in 2014 by Li et al.40, which reported significant differences in HSS between the BMSC group and CD group (DM= 8.69; 95%CI = 3.76 to 13.62; p<0.01). In 2017 Xu et al.41, found implanting BMSCs into the core decompression track resulted in better clinical outcomes than core decompression treatment by improving Harris Hip Score notably. In 2016, the systematic review by Papakostidis et al.42 found that the functional and clinical results in the reviewed trials were not consistent with a summarized estimate of effect size for the clinical outcome measures used, and the authors decided to describe the results.

**Total Hip Replacement Rate**

Meta-analysis for THR included seven RCTs16,27-32 (excluding the 25-year follow up study by Hernigou et al.32 due to differences in timepoint). The risk of THR in patients treated with CD+BMSC was 33.4% lower than the risk of patients treated with CD only (p=0.206). Though this change was not statistically significant, it is favorable to CD+BMSC and in agreement with findings of other systematic reviews done on this subject. In 2016, the systematic review by Papakostidis et al.42 found that the overall estimate of conversion to THR was in favor of cell therapy, although only marginally significant (Odds ratio [OR]=0.30, 95% CI = 0.08 to 1.06; p=0.06). In 2017, the systematic review by Xu et al.41 concluded that BMSC implantation with CD reduced the need for THR. Wang et
al.39 found the difference between CD+BMSCs and CD alone to be significant (RR=0.39; 95% CI = 0.19 to 0.78; p=0.007), with regards to the number of hips undergoing THR.

**Pain intensity (Visual Analog Scale)**

A meta-analysis of VAS scores (0-100 scale) included five trials16,27,28,30,32. VAS average improved from baseline approximately -7.4 points more in the intervention group than the CD group (95% CI = -12.113 to -2.615; p = 0.002) at 24 months. This is in agreement with a systematic review by Wang et al.39 which found a significant difference for VAS scores between the treatment and control groups at 6 months (WMD=7.08; 95%CI = -10.68 to -3.49; p<0.001), at 12 months (WMD=-7.28, 95%CI [-10.16, -4.39], p=0.000), and at 24 months (WMD= -7.93, 95%CI = -14.99 to -0.87; p=0.028). Though statistically significant, the minimum clinically significant difference in VAS, according to Kelly 2001 41 is 12 units on a scale of 100 (95% CI and 9 to 15 units). The improvement of -7.4 units on a scale of 0 through 100, shown in our review, may not be clinically significant.

**Total WOMAC score**

A meta-analysis of total WOMAC scores included two trials27,30. WOMAC improved (decreased) in average approximately -13 points more in the intervention group than the CD group (DM = - 13.456; 95% CI = -15.838 to -11.075; p<0.001). This is in agreement with the systematic review by Wang et al.39 which found a significant difference in the reduction in WOMAC scores in BMSC + CD compared to that of CD alone (WMD=10.56; 95%CI = -15.84 to -5.28; p<0.001). According to Hmamouchi et al.44 a 16.0% decrease in the total WOMAC score from baseline, was suggested to be used as a benchmark in analyzing clinical care and trials because it correlated with the greatest improvement in the transition scale category. The average percent decrease from the baseline in the WOMAC score calculated in this review (DM = - 13.456) based on reported data27,30 is equivalent to a -58% from baseline, which might be significant.

**Collapse of the femoral head**

The number of hips with collapse of the femoral head was only reported in two trials at 24 months and 25 years16,32. The results are inconclusive due to the vast difference in the time period for the two trials and further RCTs are needed. In 2016, the systematic review by Papakostidis et al.45 found the likelihood of the femoral head advancing to the collapse stage were decreased when compared to the CD group (OR=0.2; 95%CI= 0.08 to 0.6; p=0.02). However, this systematic review included three retrospective case-control trials46-48, along with one prospective pilot trial29, which were all excluded from our systematic review due to our inclusion/exclusion criteria of including only prospective RCTs. The systematic review by Xu et al.41 also found that the BMSC implantation with CD reduced the area of necrosis in the femoral head and postponed disease advancement to the stage of collapse.

**Adverse effects**

In this review, only Hauzeur et al.16 reported adverse postoperative effects (nausea, fever < 24 hours with negative bacteriological investigations, pain at the great trochanter and iliac crest)46. Accordingly, in the included trials both interventions appear safe.

**The overall completeness and applicability of evidence**

The electronic databases searched were Cochrane library, Medline via PubMed, Web of Science and EMBASE up to 1/22/2020 limited to English published articles. To find additional RCTs we hand-searched the reference sections of the included studies and reviews. The results of this systematic review are applicable to 18 to 55 years old patients with ARCO I-III ONFH and Ficat I-II. Since the reported outcomes ranged from 3 months to 25 years, this review might apply to the long-term efficacy of BMSC. Only one study had an almost equal number of male and female patients, but the rest of the studies had 2-3 times more male than females. The results might be biased toward males. These results are applicable to European countries (Belgium, France, Germany) and Asia (China, India, and Iran); they may not apply to other regions of the world.

**Heterogeneity factors in this review**

The eight RCTs included in this review had clinical and methodological heterogeneity in the diagnosis of the ONFH stage of hip disease, interventions, and outcomes. It is unclear how these parameters could affect the results as the small sample size prevented subgroup analyses for each of these variables.

**Bilateral cases**

The number of patients with bilateral hips affected varied from 1-125 hips. In one study, all study participants had bilateral ONFH, and both hips were included in the study using the hip with the larger ONFH lesion as the tested hip and the other hip with the smaller ONFH lesion as the control hip31,32. This non-randomization of hips to CD+BMSC or CD might introduce bias.

**Diagnosis**

Seven studies used ARCO (Association of Research Circulation Osseous) diagnosis system and included hips with ARCO stage I-III16,28,32. Among the seven studies, one also used the Steinberg classification system in study design and included hips with Steinberg stages 1 and II31,32. Only one study used Ficat classification system and included hips with Ficat scores I-III27.

Outcomes

Clinical outcomes included VAS scale, a self-assessment measure of pain, WOMAC a self-assessment measure of pain, stiffness and physical function outcomes, HHS, a clinical assessment measure of pain, function of daily activities and gait, absence of deformity and range of motion, and Lequesne's functional index, a self-assessment measure of pain, maximum distance walked and activities of daily living. The radiographic outcome included necrotic volume and collapse of the femoral head and THR. Due to the heterogeneity of the outcomes reported, subgroup analyses were undertaken for HHS, VAS, WOMAC and THR. No meta-analysis could be undertaken for necrotic volume (heterogeneity of the outcomes reported) and the collapse of the femoral head due to the heterogeneity of the outcomes and the very different time points.

Interventions

Methodological heterogeneity among interventions included anesthetic used for surgery (general or epidural), bone marrow harvest site (the anterior or posterior iliac crest), the terminology used for the intervention (BMAC, BBC, IMC, BMNC, and BMMSC), bone graft or not, the volume of bone marrow harvested (10-50mL), concentration and volume of cells, bone marrow insertion procedure, the seal for trochanter's access site, postoperative instructions and follow-up period. There were two different anesthetic procedures used for hip surgery, including general anesthetic and spinal epidural. Except for the epidural patients being awake during the surgery and the surgeon inadvertently making patients aware of the type of surgery, this difference has a low probability of affecting the long-term outcomes.

Seven trials used BMSC harvested from either the anterior or posterior iliac crest, and one used the subtrochanteric region. Pierini 2013 reported, “The yield of colony-founding connective-tissue progenitors was 1.6 times greater in the posterior compared with the anterior iliac crest.”

Included trials used a different procedure or protocol to concentrate the BMSCs, which are referred to as mononuclear cells (one nucleus cell vs. cells with lobes or different nuclear morphologies to be referred to as polymorphonuclear cells). Trials referred to the concentrated bone marrow stem cells with different terms: BMAC - Bone Marrow Aspirate Concentrate, BBC - Bone Based Culture, IMC - isolated mononuclear cells, BMNC - bone marrow mononuclear cells, and BMMSC - bone marrow-derived mesenchymal stem cells. All of these terms refer to the mononuclear cells concentrated from the aspirated bone marrow, and we elected to use the term and acronym Bone Marrow Stem Cells (BMSC).

In Ma et al., authors harvested a cylinder of bone from the femoral head or neck area; in the CD+BMSC group only, the buffy coat layer of centrifuged bone marrow aspirate was applied drop by drop to the surface of the bone cylinder; the concentrated bone marrow cells were applied to this buffy coat surface, and this bone cylinder was placed into the CD site. The control group patients only had the bone cylinder without a buffy coat placed into the CD site.

The time required to concentrate BMSCs after the bone marrow is harvested varied from 6 minutes to 2 weeks. Zhao et al. used one session to harvest the bone marrow followed by two weeks to increase the number of MSCs before the surgery for the CD with or without insertion of the MSCs. The volume and concentration of mononuclear cells varied widely among trials (10mL to 50mL).

Post-operatively for hip surgery, there is a period when the patient cannot place their body weight on the surgical hip, the non-weight bearing, or toe-touch weight-bearing period. For the 8 included trials, this period ranged from 10 days to 42 days (6 weeks). Rath et al.51 surveyed orthopedic surgeons for the period of postoperative weight-bearing limitation practices for high volume hip arthroscopy and found the protocol for weight-bearing was determined initially by the orthopedic surgeon, as weight-bearing as tolerated (WBAT). The weight-bearing protocol being determined empirically for each patient depending on the orthopedic surgeon or physiotherapist introduces a significant bias in determining which postoperative protocol is most advantageous. Follow-up periods varied from 2 months to 25 years.

Analysis of the influence of risk of bias on the results

The overall risk of bias of all the included RCTs was found to be either unclear or high. Among individual domains, the risk of bias for blinding was unclear in four RCTs and high in the other four. A recent systematic review, assessing clinical trials with bias due to lack of patient blinding, found that non-blinded patients could cause exaggeration of the estimated effect by up to 112%, especially with patient-reported outcomes. Patients aware of getting a placebo may lose interest during the trial or be lost to follow-up, causing dissimilar rates of attrition. Without well-planned blinding, there could be contamination, and patients in the control group may receive unintended intervention and vice versa. Apart from the lack of blinding, inadequate allocation concealment and the inadequate generation of a randomized sequence can both allow selection bias. In conclusion, the accumulated high risk of bias could greatly reduce confidence that the results represent the true effect of the treatment under investigation.

The confidence in research outcomes can be improved by the elimination of bias in various domains. To eliminate...
Implications for clinical practice

The quality of the evidence and implications for research

Quality of evidence was low for HHS scores (at 12 & 24 months), change in WOMAC scores at 24 months, and change in VAS scores due to high risk of bias and imprecision (total sample size < 400). The quality of evidence was moderate for THR, with more than 400 total participants and eight RCTs (Table 8). Future RCTs need larger sample sizes and a lower risk of bias. In order to reduce the risk of bias, researchers should clearly and prospectively outline their protocols, publish their methods of randomization, allocation concealment (i.e., using sealed envelopes to assure effective blinding of the clinicians, participants, and analyzers). Using CD plus sham skin incision/saline injections in the control group could also minimize the risk of bias for lack of blinding. It should also be clear if researchers used any public or private funding for RCTs or implemented any co-interventions, affecting the results. Balancing the intervention and the control group’s baseline characteristics also could add to the reliability of the results of the trials.

Implications for clinical practice

In the United States, the only FDA-approved stem cell-based products are hematopoietic progenitor cells derived from cord blood and these products are approved for limited use in patients with diseases of the hematopoietic system. However, there are hundreds of major clinical trials currently underway globally, focusing on more than a hundred proposed therapeutic uses with hematopoietic stem cells collected from cord blood, including treating heart disease, spinal cord injuries, liver disease, lung cancer, stroke, etc.

The main apprehension associated with stem cell therapy is the undesirable differentiation of the transplanted MSCs and their prospective ability to subdue anti-tumor immune response and initiate neoangiogenesis that may encourage tumor growth and metastasis. Consequently, the focus of future research should be on a clear description of various component aspects and signaling pathways controlling the behavior of MSCs after their in-vivo administration.

Our meta-analyses found that the addition of BMSC to CD improves the clinical outcomes in the treatment of ONFH as compared to CD, and we propose further research is needed to confirm the results. Although our systematic review found no reported adverse postoperative effects associated with BMSC addition with CD, we also propose it is important to ascertain with future research the safety profile of stem cell therapy due to the aforementioned concerns.

CONCLUSION

A total of eight RCTs were included in this systematic review. Five RCTs were assessed at high risk of bias and three at unclear risk of bias. The meta-analyses showed a statistically significant improvement from baseline of clinical outcomes, HHS (at 12 and 24 months), changes in VAS, and WOMAC scores in the BMSC + CD group with low quality of the evidence. Though statistically significant, we did not find the improvement in HHS and VAS to be clinically significant based on prior definitions of “minimum clinically significant difference”. We also found a non-statistically significant improvement of THR (after excluding the 25-year follow-up study by Hernigou et al.) with BMSC+CD and a clinically significant average percent decrease in WOMAC scores at 24 months. More high-quality RCTs with low risk of bias and large sample size will be needed to determine better the efficacy of using BMSC with CD as compared to CD alone. It would be beneficial to conduct further RCTs with more standardized and homogenous processes, as far as the CD procedure, the extraction process and the volume of BMSC, the number of BMSCs, and the postoperative care for the patients.

ACKNOWLEDGMENTS

Conflicts of interest

The authors of this review had no conflicts of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
REFERENCES


