Effectiveness of mesenchymal stem cells for treating patients with knee osteoarthritis: a meta-analysis toward the establishment of effective regenerative rehabilitation

Hirotaka Iijima1,2,3, Takuya Isho2,4, Hiroshi Kuroki2, Masaki Takahashi1 and Tomoki Aoyama2

This systematic review with a meta-analysis aimed to summarize the current evidence of the effectiveness of mesenchymal stem cell (MSC) treatment for knee osteoarthritis (OA) and to examine whether rehabilitation is an effect modifier of the effect estimate of MSC treatment. A literature search yielded 659 studies, of which 35 studies met the inclusion criteria (n = 2385 patients; mean age: 36.0–74.5 years). The meta-analysis results suggested that MSC treatment through intra-articular injection or arthroscopic implantation significantly improved knee pain (standardized mean difference [SMD]: −1.45, 95% confidence interval [CI]: −1.94, −0.96), self-reported physical function (SMD: 1.50, 95% CI: 1.09, 1.92), and cartilage quality (SMD: −1.99; 95% CI: −3.51, −0.47). However, the MSC treatment efficacy on cartilage volume was limited (SMD: 0.49; 95% CI: −0.19, 1.16). Minor adverse events (knee pain or swelling) were reported with a wide-ranging prevalence of 2–60%; however, no severe adverse events occurred. The evidence for these outcomes was “very low” to “low” according to the Grades of Recommendation, Assessment, Development and Evaluation system because of the poor study design, high risk of bias, large heterogeneity, and wide 95% CI of the effects estimate. Performing rehabilitation was significantly associated with better SMD for self-reported physical function (regression coefficient: 0.881, 95% CI: 0.049, 1.712; P = 0.039). We suggest that more high quality randomized controlled trials with consideration of the potential rehabilitation-driven clinical benefit would be needed to facilitate the foundation of effective MSC treatment and regenerative rehabilitation for patients with knee OA.

npj Regenerative Medicine (2018) 3:15; doi:10.1038/s41536-018-0041-8

INTRODUCTION

Osteoarthritis (OA) is the most common form of arthritis.1 OA ultimately results in cartilage degeneration, chronic knee pain, and disability. In 2010, knee OA was the 11th leading cause of disability worldwide, with increasing incidence over the last 2 decades.2 Current treatments have little impact on the progressive disability worldwide, with increasing incidence over the last 2 decades.2

Mesenchymal stem cells (MSCs) have emerged as a cell type with great potential for cell-based articular cartilage repair in patients with knee OA.3 Clinical trials that investigate the effects of MSC treatments in patients with knee OA have recently begun emerging,4 and results of clinical studies are continuously reported.5-10 Several meta-analyses summarize the effects of MSC treatment in patients with knee OA;5-10 these studies contribute to the establishment of effective cell-based therapies for degenerative cartilage disease. However, some of these systematic reviews included patients with focal cartilage lesions6-10 or focused on pain and physical function as treatment outcomes,5,7,9,10 with a large heterogeneity and lack of evaluation of bias risk.7-9 As knee pain would be discordant with articular cartilage status, understanding the effects of MSC treatment against OA joint degeneration and exploring the mechanisms underlying symptom-modifying MSC treatment are important. In addition, confidence in the effects estimate from meta-analysis depends on the quality of the included studies and analytical process,11 as the former can be evaluated using the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach.12 However, no meta-analysis has examined the effects of MSCs on knee OA considering the GRADE approach.

Physical factors such as rehabilitation programs are potential effect modifiers that were not well addressed in previous meta-analyses.5-10 Physical factors regulate MSC differentiation and tissue development, pointing to a potential therapeutic strategy for enhancing the MSCs injected or implanted into the knee joint.13,14 Such as the recently proposed new field “regenerative rehabilitation”.15 Regenerative rehabilitation is defined as the integration of principles and approaches from the fields of rehabilitation science and regenerative medicine.16 The efficacy of regenerative medicine may be enhanced when coupled with mechanical input. Weight-bearing might influence the structural outcome in the postoperative phase of autologous chondrocyte implantation in adults with cartilage defects.17,18 Thus, further investigation of the effects of MSC treatment in patients with knee OA and the potential role of rehabilitation (i.e., regenerative rehabilitation) as an effect modifier would be of interest.

1Department of System Design Engineering, Keio University, Yokohama, Japan; 2Department of Physical Therapy, Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan; 3Japan Society for the Promotion of Science, Tokyo, Japan and 4Rehabilitation Center, Fujioka General Hospital, Gunma, Japan

Correspondence: Hirotaka Iijima (hirotaka.4m@yts.sd.keio.ac.jp)

Hirotaka Iijima and Takuya Isho contributed equally to this work.

Received: 26 August 2017 Revised: 4 January 2018 Accepted: 5 January 2018

Published online: 17 September 2018
Potential adverse effects have a considerable impact on patient adherence to MSC treatment. To achieve a balanced perspective, a systematic review should consider the aspects of adverse events relevant to MSC treatment. Randomized controlled trials (RCTs) would be insufficient to provide evidence of benefits and harms; thus, non-RCT, such as prospective cohort studies with long-term follow-up periods should be included. However, no systematic reviews have investigated adverse events after MSC treatment, even though previous systematic reviews included both RCTs and non-RCTs. Thus, the purpose of this systematic review was (i) to examine the literature on the effects of MSCs in patients with knee OA in the clinical setting and to summarize the current evidence for their potential benefits and harms, and (ii) to examine whether rehabilitation is an effect modifier of effect estimate of MSC treatment. This study would provide a framework for a future high-quality study with the aim of developing effective cell-based regenerative rehabilitation in patients with knee OA.

RESULTS
eFigure 1 shows a flow chart of the study selection. The database search yielded 659 studies, of which 31 met the eligibility criteria. With the citation index, 4 additional studies were found in accordance with the pre-specified inclusion criteria provided in eMethod 1; in total, 35 studies were used in the meta-analysis.

Study characteristics
Table 1 shows the characteristics of the included studies. Of 35 studies, 21 (60.0%) had a single-arm prospective design, 7 (20.0%) had a quasi-experimental design, and the remaining 7 (20.0%) were RCTs. From the 35 studies, 2385 patients treated with MSC therapy were included. The mean age across 35 articles was 56.7 ± 6.78 years (36.0–74.5 years). In the 30 studies that reported sex (n = 1975 patients), 1119 patients (56.7%) were female. Twenty-nine studies (82.9%) reported the radiographic severity of knee OA (i.e., Kellgren/Lawrence [K/L] grade); however, the eligibility criteria, disease severity differed between studies. The final follow-up period was 3–60 months. Fourteen studies (40.0%) reported funding sources (eTable 1). Of the 35 studies, 25 (73.5%) used autologous and allogeneic MSC intra-articular injection, respectively. The other studies used arthroscopic autologous MSC implantation, or a combination of these procedures with high tibial osteotomy.

The rehabilitation program included patients’ education in the pre-MSC treatment phase, gradual increase in weight-bearing using crutches, use of physical therapy modalities, range of motion exercise, and muscle strength exercise (eTable 2). Notably, none of the included studies stratified for the presence of rehabilitation.

Risk of bias within studies
A summary of the Downs and Black scale for assessing bias risk is shown in eTable 3. The mean score for all 35 studies was 6.1 ± 2.1 (range, 3–12); 5.5 ± 1.6 for single-arm prospective studies; 6.3 ± 1.0 for quasi-experimental studies; and 7.9 ± 3.2 for RCT. Only two studies received a score of 1, for blinding of participants and assessors who measured key outcomes and concealed randomization of patients. The main differences between RCTs and non-RCTs included the reporting of patients’ recruitment and adequate adjustment for confounders, which is important for assessing the external and internal validities of studies.

Outcome measures
Self-reported knee pain. Nineteen studies with 27 data sets (n = 318) reported MSC treatment effects on knee pain by using the visual analog scale (VAS) pain score (Fig. 1). The mean follow-up period in these studies was 14.0 ± 12.9 months. The baseline VAS pain score in these studies was 60.2 ± 13.8 mm. Considering all 19 studies, the pooled standardized mean difference (SMD) on the VAS knee pain was −1.45 (95% confidence interval [CI]: −1.94, −0.96; P < 0.001). This statistical value implies a mean difference of 27.6 mm (95% CI: 13.4, 41.9 mm). However, effects estimates were highly heterogeneous among studies (I² = 84%). Stratification for donor type (i.e., autologous vs. allogeneic) did not much improve the heterogeneity, but the pooled SMD in autologous MSC was likely to have a larger pain relief effects than those in allogeneic MSC. A meta-regression analysis indicated that a higher score of the Downs and Black scale (i.e., low risk of bias) is significantly associated with a higher (i.e., lower effect) SMD (eTable 4). Among the subitems of the Down and Black scale and SMD, clear patients’ recruitment site was significantly associated with a higher SMD (eTable 5). Rehabilitation (i.e., using physical therapy modalities, range of motion exercise, or muscle strength exercise at least one time) was not an effect modifier of SMD (regression coefficient: 0.451, 95% CI: −1.909, 2.811; P = 0.696). Small-study effects were visually observed by two independent reviewers (eFigure 2), and the Egger’s regression test was positive for significant evidence of publication bias (P = 0.016). By using the trim-and-fill method, the adjusted SMD was −0.93 (95% CI: −1.29, −0.56; P < 0.001).

To address the possibility that effect estimates on VAS pain score and heterogeneity change if only RCTs were included in the meta-analysis, we performed a sensitivity analysis (Fig. 2). Three RCT studies with 7 data sets (n = 75) were included, and the follow-up period of all these studies was 12.0 months. The baseline VAS pain score of these studies was 60.4 ± 9.2 mm. Including only RCTs attenuated the pain relief effects (pooled SMD: −0.67, 95% CI: −1.28, −0.05; P = 0.030). This statistical value implies a mean difference of 18.1 mm (95% CI: 1.35, 34.8 mm). However, effects estimates were still highly heterogeneous among the studies (I² = 68%). Stratification for donor type slightly improved the heterogeneity, and the pooled SMD in autologous MSC was likely to have larger pain relief effects than those in allogeneic MSC. A meta-regression analysis indicated that a higher score in the Downs and Black scale and younger age were significantly associated with higher (i.e., lower effect) SMDs (eTable 6), and blinding of participants and assessors, valid outcome measures, and concealed allocation were significantly associated with higher SMDs (eTable 7). As all the included RCTs did not report a rehabilitation program, the regression coefficient could not be calculated. No small-study effect was visually observed by two independent reviewers (eFigure 3).

Self-reported physical function. Nineteen studies with 29 data sets (n = 528) reported MSC treatment effects on self-reported physical function by using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) functional, International Knee Documentation Committee (IKDC), and Lysholm scores (Fig. 3). The mean follow-up period in these studies was 17.0 ± 10.8 months. Considering all 19 studies, the pooled SMD on the self-reported physical function was 1.50 (95% CI: 1.09, 1.92; P < 0.001). This statistical value implies a mean difference of 14.7 mm (95% CI: 13.5, 15.9 mm). However, effects estimates were still highly heterogeneous among the studies (I² = 68%). Stratification for donor type slightly improved the heterogeneity, and the pooled SMD in autologous MSC was likely to have larger pain relief effects than those in allogeneic MSC. A meta-regression analysis indicated that a higher score in the Downs and Black scale and younger age were significantly associated with higher (i.e., lower effect) SMDs (eTable 6), and blinding of participants and assessors, valid outcome measures, and concealed allocation were significantly associated with higher SMDs (eTable 7). As all the included RCTs did not report a rehabilitation program, the regression coefficient could not be calculated. No small-study effect was visually observed by two independent reviewers (eFigure 3).
**Summary of included studies**

| Author | Subject population | KL grade | Treatment | Donor | Outcomes | Follow-up | Funding |
|--------|-------------------|----------|-----------|-------|----------|-----------|---------|
| Single-arm, prospective follow-up studies |
| Bui 2014**20** (Vietnam) | N = 21 | II–III | SVF injection + PRP | Auto | Lysholm score, VAS pain, MRI | 1, 3, 6 M | X |
| Centeno 2008a**21** (United States) | N = 1 (age: 36 y; M) | II | BD-MSC injection (4.56 x 10^7 cells) | Auto | VAS pain, MRI (cartilage and meniscus volumes) | 1, 3 M | – |
| Centeno 2008b**22** (United States) | N = 1 (46 y; M) | III | BD-MSC injection (2.24 x 10^7 cells) | Auto | VAS pain, functional rating index, ROM, MRI evaluation (cartilage and meniscus volumes) | 1, 3, 6 M | – |
| Davatchi 2011**23** (Iran) | N = 7 (age: 36 y; M) | II | BD-MSC injection (4.56 x 10^7 cells) | Auto | VAS pain, MRI (cartilage and meniscus volumes) | 1, 3 M | X |
| Centeno 2008b**22** (United States) | N = 1 (46 y; M) | III | BD-MSC injection (2.24 x 10^7 cells) | Auto | VAS pain, functional rating index, ROM, MRI evaluation (cartilage and meniscus volumes) | 1, 3, 6 M | – |
| Davatchi 2016**24** (Iran) | N = 4 (age: 57.8 ± 5.0 y; 50% F) | II–III | BD-MSC injection (8–9 x 10^6 cells) | Auto | VAS pain, ROM | 6 M | X |
| Emadedin 2012**25** (Iran) | N = 4 (age: 57.8 ± 5.0 y; 50% F) | II–III | BD-MSC injection (8–9 x 10^6 cells) | Auto | VAS pain, ROM | 6 M | X |
| Emadedin 2015**26** (Iran) | N = 6 (age: 70.3 ± 17.3 y; 83.3% F) | II–III | BD-MSC injection (2.0–2.4 x 10^7 cells) | Auto | VAS pain, WOMAC, MRI evaluation | 2, 6, 12, 30 M | X |
| Fodor 2016**27** (United States) | N = 6 patients 8 knees (age: 59.0 ± 7.3 y; 83.3% F) | I (N = 2) II (N = 2) III (N = 4) | SVF injection | Auto | VAS pain, WOMAC, ROM, TUG, MRI evaluation | 3, 12 M | X |
| Kim 2015c**28** (Korea) | N = 49 patients, 55 knees (age: 58.1 ± 8.9 y; 52.7% F) | I–II | AD-MSC implantation (4.3 x 10^6 cells) + AD | Auto | IKDC, Tegner activity scale | 26.7 M | – |
| Kim 2016**29** (Korea) | N = 20 patients, 24 knees (age: 57.9 ± 5.9 y; 45.0% F) | I–II | AD-MSC implantation (4.4 x 10^6 cells) + AD | Auto | IKDC, Tegner activity scale, MRI evaluation (MOCART and MOAKS) | 27.9 M | – |
| Koh 2013**30** (Korea) | N = 18 (age: 54.6 ± 7.8 y; 66.7% F) | III–IV | AD-MSC injection (1.8 x 10^6 cells) + PRP | Auto | WOMAC, Lysholm score, VAS pain, MRI evaluation (WORMS) | 24.3 M | – |
| Koh 2014a**31** (Korea) | N = 35 patients, 37 knees (age: 57.4 ± 5.7 y; 60.0% F) | I–II | AD-MSC implantation (3.8 x 10^6 cells) + AD | Auto | IKDC, Tegner activity scale, arthroscopic evaluation (ICRS grade) | 26.5 M | – |
| Koh 2015**32** (Korea) | N = 30 (age: 70.3 [65–80] y; 83.3% F) | II–III | AD-SVF (4.2 x 10^7 cells) injection + PRP + AD | Auto | Lysholm, KOOS, VAS pain, K/L grade, arthroscopic evaluation | 3, 12, 24 M | X |
| Michalek 2015**33** (Czech Republic) | N = 1114 (age: 62.0 [19–94] y; 47.8% F) | II–IV | AD-SVF injection (1.6 x 10^6 cells) + PRP | Auto | Modified KOOS, X-ray, MRI evaluation | 17.2 M | X |
| Orozco 2013**34** (Spain) | N = 12 (age: 49.0 ± 17.3 y; 50.0% F) | II (N = 4) III (N = 3) IV (N = 5) | BD-MSC injection (4.0 x 10^7 cells) | Auto | VAS pain, Lequesne index, WOMAC, PCI, SF-36 | 3, 6, 12 M | – |
| Orozco 2014**35** (Spain) | N = 12 (age: 49.0 ± 17.3 y; 50.0% F) | II–IV | BD-MSC injection (4.0 x 10^7 cells) | Auto | VAS pain score, Lequesne index, WOMAC, PCI | 3, 6, 12, 24 M | – |
| Pak 2011**36** (Korea) | N = 2 (age: 74.5 ± 6.4 y; 100% F) | III–IV | AD-MSC injection + HA + PRP + CaCl_2 + dexamethasone | Auto | VAS pain, ROM, MRI evaluation | 3 M | – |
| Sampson 2016**37** (United States) | N = 125 (age: 57.0 [23–79] y; 100% F) | III–IV | BMC injection + PRP | Auto | VAS, global patients satisfaction survey | 4.8 M | – |
| Soler Rich 2015**38** (Spain) | N = 50 (age: 57.8 ± 14.1 y; 40.0% F) | II–IV | BD-MSC injection (4.0 x 10^7 cells) | Auto | VAS, Lequesne score, WOMAC, MRI evaluation (T2 mapping, PCI) | 0, 6, 12 M | – |
| Soler 2016**39** (Spain) | N = 15 (age: 51.1 ± 10.3 y; 60.0% F) | II (N = 9) III (N = 6) | BD-MSC injection (4.1 x 10^7 cells) | Auto | VAS, Lequesne score, WOMAC, MRI evaluation (T2 mapping) | 1 W, 3, 6, 12, 48 M | X |
| Trajune 2013**40** (Thailand) | N = 5 (age: 57.2 ± 1.92 y; 80.0% F) | II | AAPBSC injection + GFAP concentrate + HA + MCS | Auto | WOMAC, KOOS | 1, 6 M | X |
| Author | Subject population | KL grade | Treatment | Donor | Outcomes | Follow-up | Funding |
|--------|--------------------|----------|-----------|-------|----------|-----------|---------|
| **Quasi-experimental studies** | | | | | | | |
| Centeno 2014 | I: N = 518 (age 54.3 ± 14.1 y) C: N = 163 (age 59.9 ± 10.3 y) | I: I (N = 223) II (N = 145) III/IV (N = 102) | I: BMC injection + PRP with adipose fat graft C: BMC injection + PRP | Auto | Improvement rating scale, LEFS, NPS | 1, 3, 6, 12 | – |
| Jo 2014 | I-a: Low dose, N = 3 (age: 63.0 ± 8.6 y; 66.7% F) I-b: Mid dose, N = 3 (age: 65.0 ± 6.6 y; 100% F) I-c: High dose, N = 12 (age: 61.0 ± 6.2 y; 83.3% F) | I-a: III (N = 2) IV (N = 1) I-b: III (N = 2) IV (N = 1) I-c: III (N = 8) IV (N = 4) | AD-MSC injection (I-a: 1.0 × 10^7, I-b: 5.0 × 10^6, I-c: 1.0 × 10^7 cells) | Auto | WOMAC, VAS pain, KSS, MRI evaluation (defect size and cartilage volume), arthroscopic evaluation (ICRS grade), biopsy | 1, 2, 3, M | X |
| Kim 2015a | I: N = 17 patients, 17 knees (age: 57.7 ± 5.8 y; 52.9% F) C: N = 37 patients, 39 knees (age: 57.5 ± 9.9 y; 62.2% F) | I–II | I: AD-MSC implantation with fibrin glue (3.9 × 10^6 cells) + AD C: AD-MSC implantation (3.9 × 10^6 cells) + AD | Auto | IKDC, Tegner activity scale, arthroscopic evaluation (ICRS grade) | 2.8 M | – |
| Kim 2015b | I: N = 20 (age: 59.1 ± 3.5 y; 65.0% F) C: N = 20 (age: 59.4 ± 3.1 y; 65.0% F) | I–II | I: AD-MSC implantation (4.0 × 10^6 cells) + AD C: AD-MSC injection (4.0 × 10^6 cells) + PRP | Auto | IKDC, Tegner activity scale, arthroscopic evaluation (ICRS grade) | 2.8 M | – |
| Koh 2012 | I: N = 25 (age: 54.2 ± 9.3 y; 68.0% F) C: N = 25 (age: 54.4 ± 11.3 y; 68.0% F) | I–II | I: AD-MSC injection (1.89 × 10^6 cells) + PRP C: PRP | Auto | Lysholm, Tegner activity scale, VAS pain | 3, 16, M | – |
| Nguyen 2017 | I: N = 15 (age: 58.6 ± 6.5 y; 80.0% F) C: N = 15 (age: 58.2 ± 5.7 y; 80.0% F) | I–II | I: AD-SVF injection (1.89 × 10^6 cells) + AM + PRP C: AM + PRP | Auto | WOMAC, modified VAS pain, Lysholm, MRI | 1, 6, 12, 18 | M |
| Pers 2016 | I-a: Low dose, N = 6 (age: 63.2 ± 4.1 y; 50.0% F) I-b: Mid dose, N = 6 (age: 65.5 ± 8.1 y; 50.0% F) I-c: High dose, N = 6 (age: 65.2 ± 2.3 y; 66.7% F) | I–II | I-a: III (N = 2) IV (N = 1) I-b: III (N = 1) IV (N = 5) I-c: III (N = 0) IV (N = 6) | Auto | WOAM, Global knee pain, PGA, KOOS, SAS, SF-36, MRI evaluation | 1, 2, 3, M | X |
| **Randomized controlled trials** | | | | | | | |
| Gupta 2016 | Cohort 1: I-a (Low dose): N = 10 (age: 58.1 ± 8.2 y; 70.0% F) I-b (Mid dose): N = 10 (age: 57.3 ± 9.5 y; 80.0% F) C-a: N = 10 (age: 54.9 ± 8.3 y; 100.0% F) Cohort 2: I-c (High dose): N = 10 (age: 53.0 ± 6.7 y; 80.0% F) | I–II | I-a: II (N = 4) III (N = 6) I-b: II (N = 1) III (N = 9) C-a: II (N = 3) III (N = 7) I-c: II (N = 1) III (N = 9) I-d: II (N = 3) III (N = 7) C-b: II (N = 2) III (N = 8) | Allo | BD-MSC injection (I-a: 25 × 10^6, I-b: 50 × 10^6, I-c: 75 × 10^6 cells, I-d: 150 × 10^6 cells) + HA C: HA | 12 M | X |
| Koh 2014b | I: N = 21 (age: 54.2 ± 2.9 y; 76.2% F) C: N = 23 (age: 52.3 ± 4.9 y; 73.3% F) | I–II | I: HTO + AD-MSC implantation + PRP C: HTO + PRP | Auto | Lysholm, KOOS, VAS pain, FTA, arthroscopic evaluation (Kanamia grade) | 24.4 M | – |
| Lamo-Espinosa 2016 | I-a (Low dose): N = 10 (age: 65.9 [IQR: 59.5, 70.6] y; 60.0% F) I-b (High dose): N = 10 (age: 57.8 [IQR: 55.0, 60.8] y; 20.0% F) C: N = 10 (age: 60.3 [IQR: 55.1, 61.1] y; 30.0% F) | I–II | I-a: II (N = 7) III (N = 3) IV (N = 4) C: II (N = 4) III (N = 3) IV (N = 2) | Auto | VAS, WOMAC, ROM, X-ray, MRI (WORMS) | 3, 6, 12 M | X |
| Varma 2010 | I: N = 25 (age: 50.7 ± 5.4 y) C: N = 25 (age: 48.2 ± 5.1 y) | – | I: BMC injection + AD C: AD | Auto | VAS pain, OAOS | 1, 2, 3, 6 M | – |
**Table 1 continued**

| Author Subject population | KL grade | Treatment | Donor | Outcomes Follow-up | Funding |
|---------------------------|----------|-----------|-------|--------------------|---------|
| Vega 2015\(^{50}\) (Spain) | I: N = 15 (age: 56.6 ± 9.6 y; 60.0% F) II (N = 6) III (N = 5) IV (N = 3) | Allo | Hospital for special surgery knee-rating | 6.12.24 M | - |
| Wakitani 2002\(^{16}\) (Japan) | N = 24 (age: 49.0 [24–54] y; 54.0% F) | Auto | HTO + BD-MSC implantation | 50.0% F | - |
| Wong 2013\(^{12}\) (Singapore) | N = 28 (age: 49.0 [24–54] y; 54.0% F) | Auto | HTO + BD-MSC implantation | 50.0% F | - |

MRI findings in articular cartilage. Two studies with 4 data sets (n = 20) reported the MSC treatment effect on cartilage volume, evaluated using magnetic resonance imaging (MRI; Fig. 5a). The mean follow-up period of these studies was 5.3 ± 1.5 months. In these analyses, two single case reports from the same authors\(^{21,22}\) were combined, as these case reports included patients with a similar clinical status. The pooled SMD on the cartilage volume was 0.49 (95% CI: −0.19, 1.16; P = 0.160), a non-significant small effect size. Excluding the combined two case reports resulted in similar results (pooled SMD: 0.51, 95% CI: −0.23, 1.26; P = 0.180).

The 5 other studies with 7 data sets (n = 95) reported MSC treatment effects on cartilage quality by using the poor cartilage index (PCI), dGEMERIC index, and T2 mapping values, evaluated using MRI (Fig. 5b). The mean follow-up period in these studies was 16.3 ± 15.4 months. The pooled SMD on the cartilage quality was −1.99 (95% CI: −3.51, −0.47; P < 0.001), a significantly heightened effect size (SMD ≥ 0.8), with high heterogeneity (I\(^2\) = 91%). When the pooled SMD was evaluated in each outcome measure, it became higher in the PCI but became insignificant in the dGEMERIC index, and heterogeneity improved markedly. A meta-regression analysis indicated that the presence of funding source was a significant factor associated with a higher (i.e., lower effect) SMD (eTable 10). No small-study effect was visually observed from funnel plots by two independent reviewers (efigures 6 and 7).

Adverse events

Of 35 studies, 17 (48.6%) reported adverse events related to MSC treatment. Adverse events included knee pain or swelling. eFigure 8 summarizes the event rates with their 95% CIs. Owing to the large clinical and statistical heterogeneity among the studies, we did not pool the adverse event rates. In 10 studies that reported timing of adverse events, 5,6,31,32,34,37–39,45,50 knee pain or swelling occurred within 1 week after MSC treatment; these symptoms were treatable with pain medication.

Summary of quality of evidence

Table 2 shows a summary of evidence according to the GRADE approach.\(^{12}\) The effects estimate was downgraded in all outcome measures. None of these effects estimates were upgraded. Each meta-analysis scored 1 (very low) or 2 (low) with the GRADE approach, indicating very little (i.e., the true effect is likely to be substantially different from the effect estimate) or limited (i.e., the true effect may be substantially different from the effect estimate) confidences of the effects estimate.\(^{12}\)
Estimate. Detail information about rehabilitation was lacking, effect is likely to be substantially different from the effects if we only included RCTs in the meta-analysis; therefore, the true analyses showed that these GRADE ratings were comparable even approach because of the poor study design, high risk of bias, large comprehensive assessment of relevant articles by adding new needed to facilitate a foundation of effective MSC therapy and more high quality RCTs with strati
treatment on self-reported physical function. We suggest that but rehabilitation was a signi
cant tendency to improve cartilage volume, MSC evidence of both benefits and harms of MSCs treatment on degenerative knee OA with a longer follow-up period that cannot be adequately determined by reviewing only RCTs. The mean differences after intervention were exceeding the minimum for clinically important differences, and meta-analysis further added 28 non-RCTs and 4 RCTs to the previous meta-analysis, which enable us to examine the latest evidence of both benefits and harms of MSCs treatment on degenerative knee OA with a longer follow-up period that cannot be adequately determined by reviewing only RCTs. We found that the pooled effect size on the VAS pain score exceeded the effects of nonsteroidal anti-inflammatory drugs and corticosteroid injections, consistent with previous meta-analyses. The mean differences after intervention were ≥10% for both pain and self-reported physical function, exceeding the minimum for clinically important differences, and meeting the responder criteria of the Outcome Measures in Rheumatology Clinical Trials and Osteoarthritis Research Society International. However, we found a large heterogeneity among studies, which was partly explained by the level of risk of bias, cell donor type, and study design. Including only RCTs, which has a lower risk of bias than non-RCTs, in the meta-analysis attenuated the effects of MSC treatment in improving knee pain and self-reported physical function, supporting this interpretation. The observed effects from RCTs had a wide 95% CI, and clinical action would differ if the true SMD was the upper or lower boundary of the 95% CI. This suggests the need for a larger number of RCTs to elucidate whether MSC treatment can provide clinical benefit to patients with knee OA.

**DISCUSSION**

This systematic review and meta-analysis found that MSC treatment significantly improved knee pain and self-reported physical function in patients with knee OA. While MSC treatment has an insignificant tendency to improve cartilage volume, MSC treatment significantly improved cartilage quality. However, these data should be interpreted with caution because the quality of evidence was "very low" to "low" according to the GRADE approach because of the poor study design, high risk of bias, large heterogeneity, and wide 95% CI of the pooled SMD. Sensitivity analyses showed that these GRADE ratings were comparable even if we only included RCTs in the meta-analysis; therefore, the true effect is likely to be substantially different from the effects estimate. Detail information about rehabilitation was lacking, but rehabilitation was a significant effect modifier of MSC treatment on self-reported physical function. We suggest that more high quality RCTs with stratification for rehabilitation are needed to facilitate a foundation of effective MSC therapy and regenerative rehabilitation.

The search strategies used in this study provide a more comprehensive assessment of relevant articles by adding new findings to the recent meta-analysis for the clinical efficacy of MSCs transplantation for knee OA and focal cartilage defect up to a maximum 24 months follow-up. Indeed, the current meta-analysis further added 28 non-RCTs and 4 RCTs to the previous meta-analysis, which enable us to examine the latest evidence of both benefits and harms of MSCs treatment on degenerative knee OA with a longer follow-up period that cannot be adequately determined by reviewing only RCTs.

We found that the pooled effect size on the VAS pain score exceeded the effects of nonsteroidal anti-inflammatory drugs and corticosteroid injections, consistent with previous meta-analyses. The mean differences after intervention were ≥10% for both pain and self-reported physical function, exceeding the minimum for clinically important differences, and meeting the responder criteria of the Outcome Measures in Rheumatology Clinical Trials and Osteoarthritis Research Society International. However, we found a large heterogeneity among studies, which was partly explained by the level of risk of bias, cell donor type, and study design. Including only RCTs, which has a lower risk of bias than non-RCTs, in the meta-analysis attenuated the effects of MSC treatment in improving knee pain and self-reported physical function, supporting this interpretation. The observed effects from RCTs had a wide 95% CI, and clinical action would differ if the true SMD was the upper or lower boundary of the 95% CI. This suggests the need for a larger number of RCTs to elucidate whether MSC treatment can provide clinical benefit to patients with knee OA.
The strength of this meta-analysis is that we estimated pooled SMD for structural outcomes of articular cartilage evaluated by MRI. This effect estimate was based on only 2 non-RCTs with 4 data sets, raising the need for high quality RCTs for examination of the structural modifying effects of MSC treatment. We found a discrepancy between MSC efficacy on cartilage quality and MSC efficacy on cartilage quantity (volume). While MSC treatment improved cartilage quality, it did not significantly improve cartilage volume. Although these results should be interpreted cautiously because the studies that evaluated cartilage quality differed from that evaluated cartilage volume, we found that MSC treatment may have a limited therapeutic effect on cartilage volume. Three of these 4 data sets were based on data from patients with severe knee OA (K/L grade ≥3), which may cause limited efficacy in improving cartilage volume. Furthermore, the mean follow-up period in these studies was within 6 months, which might be too short to show a biological effect. One high quality study found that MSC injection particularly improved knee pain when a relatively large number of MSCs was used, but a significant increase in cartilage volume did not accompany this pain reduction, indicating that improved knee pain is not necessarily attributable to increased cartilage volume. Although this meta-analysis only included outcome measures for articular cartilage, some included studies found that MSC treatment improved subchondral bone edema and meniscus thickness, which are predictors of knee pain severity. Improved knee pain after autologous chondrocyte implantation on cartilage defects moderately correlated with bone edema, but not the cartilage structure evaluated using MRI. Further studies that investigate the mechanism of pain reduction after MSC treatment in patients with knee OA would be of interest.

Physical factors regulate MSC differentiation and tissue development, pointing to a potential therapeutic strategy for enhancing the MSCs injected into the knee joint. Weight-bearing might influence the structural outcome evaluated by MRI in the postoperative phase of autologous chondrocyte implantation. The mean follow-up period after MSC treatment was 3–60 months in the included studies, which includes some rehabilitation and physical activity programs in the post-MSC treatment phase. These post-MSC rehabilitations might affect the effects of cell-based therapy. Indeed, the presence of rehabilitation was a significant effect modifier of SMD on self-reported functional. Although the presence of a rehabilitation program was not a significant effect modifier of the estimated effect on VAS pain score, rehabilitation does not necessarily have no impact; the lack of statistical power due to a small number of studies in the meta-analysis and the lack of details of rehabilitation program in each article may explain this absence. As physiological stimulation such as moderate level exercise, ultrasound irradiation, and mechanical loading after joint distraction may enhance cartilage regeneration after MSC injection in a preclinical study, applying exogenous stimulation may be one strategy for enhancing the injected MSCs. This point is particularly important because the lower boundary of the 95% CI of SMD on knee pain and physical function corresponds to the lower effect size in the meta-analysis of RCTs. As all the included RCTs did not report (perform) rehabilitation and none of the included non-RCTs stratified for rehabilitation program, investigating the effects of rehabilitation on the SMD of MSC treatment would be of interest in future studies. Rehabilitation programs was differed among the included studies; thus, this review highlights the need for a standardized rehabilitation program that encompasses at least weight-bearing schedule, range of motion exercise, and muscle strength exercise, which would influence the therapeutic effect of MSCs to facilitate further comparisons among studies. The implementation of longitudinal activity-based questionnaires might help address this question.

We observed a large heterogeneity of adverse event rates among the included studies; this observation limits our ability to summarize the adverse event rate. The causes of heterogeneity in this study are unclear. Detailed reports on adverse events are sparse, which may have contributed to the heterogeneity. Nevertheless, we found only minor adverse events (knee pain/swelling) after MSC treatment, indicating that benefits may outweigh harms of MSC treatment of knee OA. These findings can be achieved by reviewing the data from both non-RCTs and RCTs, which is the strength of the present meta-analysis. Most adverse events occurred within 1 week following MSC treatment. Conversely, pain or swelling that persists for more than 1 week should be interpreted as a rare and potentially severe adverse event that might contribute to arthrogenic muscle inhibition. Close attention to adverse events may be key to the clinical success in optimizing post-MSC treatment of knee OA.
Autologous MSCs are a widely selected source to minimize the immune response and an excellent therapeutic option for treating OA. Most included trials used autologous MSCs to eliminate immune rejection, while 2 of 35 articles attempted to investigate the potential application of allogeneic MSCs. No observed severe adverse event indicates the safety of allogeneic MSCs for applying knee OA. The present meta-analysis revealed that the therapeutic effects of VAS pain score and self-reported physical function were likely higher in autologous than in allogeneic MSCs. However, direct comparisons of the therapeutic effects between autologous and allogeneic MSCs are difficult because these are based on data from different studies.

| Study                          | n   | SMD (95% CI) | Weight, % | Self-reported Physical Function |
|-------------------------------|-----|-------------|-----------|---------------------------------|
| **Autologous**                |     |             |           |                                 |
| **WOMAC Physical Functional Score** |     |             |           |                                 |
| Emardin, 2012 (Iran)          | 6   | 1.70 (0.29, 3.10) | 2.9 |                                 |
| Jo, 2014; Low-dose (Korea)    | 3   | 0.45 (-1.21, 2.10) | 2.6 |                                 |
| Jo, 2014; Mid-dose (Korea)    | 3   | 1.17 (-0.79, 3.12) | 2.2 |                                 |
| Jo, 2014; High-dose (Korea)   | 12  | 0.91 (0.06, 1.76)  | 3.6 |                                 |
| Lamo-Espinosa, 2016; Low-dose (Spain) | 10  | 1.29 (0.31, 2.27) | 3.5 |                                 |
| Lamo-Espinosa, 2016; High-dose (Spain) | 10  | 1.05 (0.10, 2.00) | 3.5 |                                 |
| Orozo, 2013 (Spain)           | 12  | 0.77 (-0.07, 1.60) | 3.7 |                                 |
| Pers, 2016; Low-dose (France) | 6   | 1.69 (0.29, 3.09)  | 2.9 |                                 |
| Pers, 2016; Mid-dose (France) | 6   | 0.75 (-0.44, 1.94) | 3.2 |                                 |
| Pers, 2016; High-dose (France) | 6   | 0.52 (-0.64, 1.68) | 3.2 |                                 |
| Soler Rich, 2015 (Spain)      | 50  | 0.14 (-0.26, 0.53) | 4.1 |                                 |
| Soler, 2016 (Spain)           | 15  | 1.31 (0.51, 2.10)  | 3.7 |                                 |
| Turajane, 2013 (Thailand)     | 5   | 4.79 (1.82, 7.76)  | 1.3 |                                 |
| **Subtotal (Random effects model)** | 144 | 0.97 (0.56, 1.38) | 40.3 |                                 |
| **IKDC Score**                |     |             |           |                                 |
| Kim, 2015a-1 (Korea)          | 39  | 2.39 (1.80, 2.98)  | 4.0 |                                 |
| Kim, 2015a-2 (Korea)          | 17  | 2.99 (1.98, 4.00)  | 3.4 |                                 |
| Kim, 2015b-1 (Korea)          | 20  | 1.38 (0.69, 2.08)  | 3.8 |                                 |
| Kim, 2015b-2 (Korea)          | 20  | 2.74 (1.85, 3.63)  | 3.6 |                                 |
| Kim, 2015c (Korea)            | 55  | 3.65 (3.03, 4.26)  | 3.9 |                                 |
| Kim, 2016 (Korea)             | 24  | 2.94 (2.10, 3.77)  | 3.7 |                                 |
| Koh, 2014a (Korea)            | 60  | 2.54 (2.06, 3.03)  | 4.1 |                                 |
| **Subtotal (Random effects model)** | 235 | 2.65 (2.12, 3.18) | 26.4 |                                 |
| **Lysholm Score**             |     |             |           |                                 |
| Bui, 2014 (Vietnam)           | 21  | 2.13 (1.36, 2.90)  | 3.7 |                                 |
| Koh, 2012 (Korea)             | 25  | 1.68 (1.03, 2.33)  | 3.9 |                                 |
| Koh, 2013 (Korea)             | 18  | 2.54 (1.64, 3.44)  | 3.6 |                                 |
| Koh, 2015 (Korea)             | 30  | 1.36 (0.80, 1.93)  | 4.0 |                                 |
| Nguyen, 2017 (Vietnam)        | 15  | 1.10 (0.32, 1.88)  | 3.7 |                                 |
| **Subtotal (Random effects model)** | 109 | 1.71 (1.25, 2.17) | 18.9 |                                 |
| **Allogeneic**                |     |             |           |                                 |
| **WOMAC Physical Functional Score** |     |             |           |                                 |
| Gupta, 2016; Low-dose (India) | 10  | 0.65 (-0.25, 1.56) | 3.6 |                                 |
| Gupta, 2016; Mid-dose (India) | 10  | -0.17 (-1.05, 0.71) | 3.6 |                                 |
| Gupta, 2016; High-dose (India) | 10  | 0.54 (-0.35, 1.44) | 3.6 |                                 |
| Gupta, 2016; Very high-dose (India) | 10  | -0.02 (-0.90, 0.85) | 3.6 |                                 |
| **Subtotal (Random Effects Model)** | 40  | 0.24 (-0.20, 0.68) | 14.3 |                                 |
| **Overall (Random effects model)** | 528 | 1.50 (1.09, 1.92) | 100 |                                 |
| **Test for heterogeneity: Chi² = 23.91 (P = 0.020), I² = 50%** | | | | |
| **Test for heterogeneity: Chi² = 24.56 (P < 0.001), I² = 76%** | | | | |
| **Test for heterogeneity: Chi² = 8.19 (P = 0.080), I² = 51%** | | | | |
| **Test for heterogeneity: Chi² = 2.42 (P = 0.490), I² = 0%** | | | | |
| **Test for heterogeneity: Chi² = 206.35 (P < 0.001), I² = 86%** | | | | |

![Fig. 3](image-url)  
SMD and 95% CI for the self-reported physical functional outcome between pre and post MSC treatment at final follow-up. The diamond represents the pooled SMD using the DerSimonian-Laird method. The vertical line at 0 represents no difference. MSC treatment was effective in improving self-reported physical function (pooled SMD: 1.50, 95% CI: 1.09, 1.92; P < 0.001). SMDs were highly heterogeneous among studies (I²: 86%; P < 0.001)
of allogeneic MSCs were RCTs, which had lower risks of bias than those of autologous MSCs, which might have contributed to the lower therapeutic effect. Thus, direct comparison between autologous and allogeneic MSCs in the same trial would be of interest. This systematic review included patients with knee OA diagnosed either radiographically or clinically, and excluded those with a focal cartilage defect. Thus, the observed effect of MSCs on clinical outcomes may not hold true in patients with focal cartilage defects. As knees with OA have diffuse cartilage loss rather than an isolated cartilage lesion, several researchers have sought to assess the effect of inter-articular MSC injections rather than implantation to a focal lesion. Whereas MSC implantation on focal cartilage defects in both preclinical and clinical studies is effective in cartilage repair, the cartilage repair effects of intra-articular injection is controversial. We found that the type of treatment was a strong effect modifier of MSC treatment on physical function on physical function. It should be highlighted that 2 studies failed to detect a clear dose-response relationship between injected MSC and cartilage volume and cartilage quality, thereby no effects estimates were upgraded in the GRADE approach. Mamidi et al. recently suggested that investigating post-transplanted MSC behavior and how to enhance the potency of the transplanted MSCs are the major challenges to be directly solved in future research. We could not address post-injected MSC behavior in the diseased microenvironment; investigating the kinematics of injected MSCs is needed to enhance their disease-modifying effects.

The present study has some limitations. First, this meta-analysis included non-RCTs with 3 case reports. As non-RCTs would have greater bias and more confounders than RCTs, evaluating MSC efficacy using only RCTs might be preferable. Thus, we performed a sensitivity analysis and calculated the effect estimate based on RCTs. Meta-analyses that include non-RCTs can provide evidence of effects that are difficult to detect using a RCT, such as long-term effects and adverse events. Evaluating the beneficial and harmful effects of MSC treatment would be needed to make decisions about the clinical utility of MSC treatment. As discussed previously, as no RCTs have performed rehabilitation, the present meta-analysis, which included non-RCTs, could shed light on the importance of rehabilitation as a new strategy for enhancing functional improvement after MSC treatment and would set a basis for future high quality RCTs. Second, this meta-analysis included 35 studies, but few studies were available for use in the meta-analysis of structural outcomes. This dearth is attributable to the absence of a standard system for evaluating cartilage regeneration. Many studies that use MRI to evaluate cartilage regeneration are only qualitative, using validated imaging outcomes would be integral for scientifically validating cell-based therapies and precipitously advancing efficacy. Third, the pooled SMD included the effects of contervention such as PRP with injected or implanted MSC. PRP improves knee pain and physical function in patients with knee OA and has a similar effect to MSC injection; the pooled SMD might be attributed to the cointervention. Nevertheless, we confirmed that use of PRP was not a significant predictor of the pooled SMD (data not shown). Fourth, many studies included in this meta-analysis were performed by the same group of investigators. Thus, caution is required when interpreting the effect estimate, and further studies from different investigators are needed to elucidate the effects of MSCs on knee OA. Finally, a protocol for this systematic review has not been registered. However, protocol registration was not associated with outcome reporting bias in the meta-analysis, and the outcome measures were extracted according to the highest rank on the pain and functional outcome hierarchy, determined a priori.

In conclusion, MSC treatment improves knee pain, physical function, and cartilage quality, without any severe adverse events. However, evidence for these outcomes that are considered critical for clinical decision making was “very low” to “low” according to the GRADE system because of the poor study design, high risk of bias, large heterogeneity, and wide 95% CI of the effects estimate. These GRADE ratings were similar even if only high quality RCTs were included in the meta-analysis. Detail information about rehabilitation is lacking; therefore, the role of rehabilitation in MSC treatment in patients with knee OA is unclear. However, rehabilitation was a significant effect modifier of better MSC treatment on self-reported physical function, supporting a concept of the newly born field, regenerative rehabilitation. Integration of rehabilitation into MSC-based therapy may be beneficial at least in improving physical function. These findings would help researchers and clinicians in designing future high quality clinical trials.
METHODS

This study was conducted in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement, PRISMA protocols (PRISMA-P), meta-analysis of observational studies in epidemiology (MOOSE) checklist, and Cochrane handbook for systematic reviews of interventions. A detailed protocol for this systematic review has not been previously published and registered.

Literature search and study selection

The electronic databases of PubMed, Physiotherapy Evidence Database (PEDro), Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Cochrane Central Register of Controlled Trials were used. Searches used combined key terms, including “osteoarthritis, knee,” “transplantation,” “stem cells,” and “stromal cells,” using Medical Subject Headings terms. A database search strategy and determining inclusion are provided in the eMethods 1 and 2.

Outcome measures and data extraction

The primary outcomes in this review were (i) pain, (ii) self-reported physical function, (iii) structural outcomes of articular cartilage evaluated using MRI, and (iv) adverse events relevant to MSC treatment. Two reviewers independently extracted the data regarding authors, country, study design (single-arm, prospective follow-up studies, quasi-experimental studies, and RCTs), subject population, K/L grade, treatment, cell donor type, outcome measures, follow-up period, rehabilitation program, and funding sources using standardized data forms. When an article reported outcomes using multiple pain and functional scales, we used only the scale with the highest rank on the pain and functional outcome hierarchy, in accordance with previous recommendations and meta-analyses (eMethod 3).

Data analysis

Percent agreement of duplicate study removal and interrater reliability of title/abstract and full-text screening between the two reviewers were evaluated. For the meta-analysis, pooled estimates and 95% CIs for SMDs for changes in outcomes were calculated using the DerSimonian-Laird method. The SMD was calculated for paired samples using the within-patient change for patients treated with MSC divided by the pooled standard deviation (SD). Formulas for calculating the pooled SD and pooled SMD are provided in Table 1.

| Study                             | n    | SMD (95% CI)    | Weight, % |
|----------------------------------|------|----------------|-----------|
| Autologous                       |      |                |           |
| Centeno, 2008 (United States)    | 2    | 0.26 (-2.19, 2.71) | 7.6     |
| Jo, 2014; Low-dose (Korea)       | 3    | 2.55 (-0.39, 5.49) | 5.3     |
| Jo, 2014; Mid-dose (Korea)       | 3    | 0.14 (-1.47, 1.75) | 17.7    |
| Jo, 2014; High-dose (Korea)      | 12   | 0.44 (-0.37, 1.25) | 69.4    |
| Overall (Random effects model)   | 20   | 0.49 (0.19, 1.16) | 100     |
| Test for heterogeneity: \( \chi^2 = 2.21 \) (\( P = 0.550 \)), \( I^2 = 0\% \) |

Test for heterogeneity: \( \chi^2 = 3.55 \) (\( P = 0.060 \)), \( I^2 = 72\% \)

| Study                             | n    | SMD (95% CI)    | Weight, % |
|----------------------------------|------|----------------|-----------|
| Autologous                       |      |                |           |
| Centeno, 2008 (Spain)            | 12   | -2.82 (-4.01, -1.64) | 17.6    |
| Soler Rich, 2015 (Spain)         | 50   | -4.15 (-4.85, -3.44) | 18.9    |
| Subtotal (Random effects model)  | 62   | -3.57 (-4.86, -2.29) | 36.4    |
| Test for heterogeneity: \( \chi^2 = 3.55 \) (\( P = 0.060 \)), \( I^2 = 72\% \) |

Test for heterogeneity: \( \chi^2 = 0.03 \) (\( P = 0.860 \)), \( I^2 = 0\% \)

| Study                             | n    | SMD (95% CI)    | Weight, % |
|----------------------------------|------|----------------|-----------|
| Autologous                       |      |                |           |
| Centeno, 2008 (United States)    | 2    | 0.26 (-2.19, 2.71) | 7.6     |
| Jo, 2014; Low-dose (Korea)       | 3    | 2.55 (-0.39, 5.49) | 5.3     |
| Jo, 2014; Mid-dose (Korea)       | 3    | 0.14 (-1.47, 1.75) | 17.7    |
| Jo, 2014; High-dose (Korea)      | 12   | 0.44 (-0.37, 1.25) | 69.4    |
| Overall (Random effects model)   | 20   | 0.49 (0.19, 1.16) | 100     |
| Test for heterogeneity: \( \chi^2 = 2.21 \) (\( P = 0.550 \)), \( I^2 = 0\% \) |

Test for heterogeneity: \( \chi^2 = 54.84 \) (\( P < 0.001 \)), \( I^2 = 91\% \)

Fig. 5 SMD and 95% CI for cartilage volume (a) and cartilage quality (b) between pre and post MSC treatment at final follow-up. The diamond represents the pooled effect size using the DerSimonian-Laird method. The vertical line at 0 represents no difference. While MSC treatment has a non-significant tendency to improve cartilage volume (pooled SMD: 0.49, 95% CI: −0.19, 1.16; \( P = 0.160 \)), MSC treatment was effective in improving cartilage quality (pooled SMD: −1.99, 95% CI: −3.51, −0.47; \( P < 0.001 \)). SMDs for cartilage quality were highly heterogeneous among studies (\( I^2: 91\% \); \( P < 0.001 \)).
shown in eMethod 5. The meta-analyses were performed using Review Manager Version 5.3 (Nordic Cochrane Center, Cochrane Collaboration, Copenhagen, Denmark). We used a forest plot to represent the meta-analysis results in accordance with a previous study. The size of the SMD was interpreted using Cohen’s $d$ ($<0.5$: small effect size, 0.5–0.8: moderate effect size, and $\geq 0.8$: large effect size). As a clinical frame of reference, a small effect is equivalent to the effect of non-steroidal anti-inflammatory drugs on knee pain in OA trials. A moderate effect is equivalent to the effect of corticosteroid injections on knee pain. When mean and SD values were not directly reported in an article, they were calculated from other available data, if possible (eMethod 6). To test for publication bias, we used a funnel plot and Egger’s test, where publication bias is the tendency for positive trials to be published and the tendency for negative or null trials to not be published. We interpreted $P$-values of $<0.10$ to indicate the existence of publication bias, as practiced by a previous study. When studies are relatively few, the power of the test is too low to distinguish chance from real asymmetry; we tested for publication bias only when least 10 studies were included in the meta-analysis, and if present, adjustment was planned using a trim-and-fill method. As SMD would be difficult to interpret in a clinical context, the mean differences in pain and functional outcomes were also calculated and compared with minimum clinically important difference (eMethod 7). Furthermore, we performed prespecified sensitivity analyses to provide pooled SMD with 95% CI by using the data from RCTs only.

Study heterogeneity was assessed using the $I^2$ statistic and $Q$ statistic. If $I^2$ was $\geq 50$, random effects meta-regression was performed using the certain parameters selected a priori including the presence of rehabilitation, defined when patients were treated using physical therapy modalities, range of motion exercise, or muscle strength exercise at least one time after MSC treatment (eMethod 8). Adverse events were evaluated in each study, and adverse event rates were calculated from the numbers of events and sample sizes by using the Comprehensive Meta-Analysis software (Biostat, Inc., Englewood, NJ, USA). All other statistical analyses were performed using JMP Pro 12.2 (SAS Institute, Cary, NC, USA).

Additional methods
Additional methods for assessment of risk of bias and GRADE approach are provided in eMethods in the Supplement.

Data availability
Data available on request from the authors.

ACKNOWLEDGEMENTS
The authors thank members of the regenerative rehabilitation team (Kyoto University, Kyoto) for their assistance and advice. This study was supported in part by a Grant-in-Aid from the Japan Society for the Promotion of Science (https://www.jsps.go.jp/) for Research Fellows to HI.

AUTHORS CONTRIBUTIONS
All authors met following criteria: substantial contributions to the conception and design of the study, acquisition of data, or analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; final approval of the version to be submitted; and accountability for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The specific contributions of the authors are as follows: Conception and design of the study: H. I., T. I., and T. A. Analysis and interpretation of the data: H. I., T. I., H. K., and T. A. Drafting of the article: H. I., T. I., M. T., and T. A. Critical revision of the article for important intellectual content: H. I., T. I., H. K., and T. A. Final approval of the article: H. I., T. I., H. K., M. T., and T. A. Statistical expertise: H. I. and T. I. Obtaining of funding: H. I. Collection and assembly of data: H. I. and T. I.
Kwon, H. J., Lee, G. S. & Chun, H. Electrical stimulation drives chondrogenesis of bone marrow mesenchymal stem cells for the management of osteoarthritis: state of the art and perspectives. Osteoarthr. Cartil. 23, 2027–2035 (2015).

Mamidi, M. K., Das, A. K., Zakaria, Z. & Bhonde, R. Mesenchymal stromal cells for treatment of knee osteoarthritis: multicenter randomized controlled trial (Phase II/III). J. Transl. Med. 17, 343 (2015).

Centeno, C. J. et al. Regeneration of meniscus cartilage in a knee treated with autologous bone marrow mesenchymal stem cells: two-year follow-up results. Transplantation 97, e66–e68 (2016).

Koh, Y. G. et al. Mesenchymal stem cell injections improve symptoms of knee osteoarthritis. Arthroscopy 29, 748–755 (2013).

Michalek, J. et al. WITHDRAWN: autologous adipose tissue-derived stromal vascular fraction cells application in patients with osteoarthritis. Cell Transpl. (2015).

Orozco, L. et al. Treatment of knee osteoarthritis with autologous mesenchymal stem cells: a pilot study. Transplantation 95, 1535–1541 (2013).

Koh, Y. G. et al. Mesenchymal stem cell injections improve symptoms of knee osteoarthritis. Arthroscopy 29, 748–755 (2013).

Solera, R. et al. Final results of a phase I-II trial using ex vivo expanded autologous mesenchymal stromal cells for the treatment of osteoarthritis of the knee confirming safety and suggesting cartilage regeneration. Knee 23, 647–654 (2016).

Koh, Y. G. et al. Mesenchymal stem cell implantation in osteoarthritic knees: is fibrin glue effective as a scaffold? Am. J. Sports Med. 43, 176–185 (2015).

Kim, Y. S. et al. Comparative matched-pair analysis of the injection versus implantation of mesenchymal stem cells for knee osteoarthritis. Am. J. Sports Med. 43, 2738–2746 (2015).

Koh, Y. G. & Choi, J. Y. Infrapatellar fat pad-derived mesenchymal stem cell therapy for knee osteoarthritis. Knee 19, 902–907 (2012).

Nguyen, P. D. et al. Comparative clinical observation of arthroscopic microfracture in the presence and absence of a stromal vascular fraction injection for osteoarthritis. Stem Cells Transl. Med. 6, 187–195 (2017).

Gupta, P. K. et al. Efficacy and safety of adult human bone marrow-derived, cultured, pooled, allogeneic mesenchymal stromal cells (Stempeucel(R)): preclinical and clinical trial in osteoarthritis of the knee joint. Arthritis Res. Ther. 18, 301 (2016).

Koh, Y. G., Kwon, O. R., Kim, Y. S. & Choi, J. Y. Comparative outcomes of open-wedge high tibial osteotomy with platelet-rich plasma alone or in combination with mesenchymal stem cell treatment: a prospective study. Arthroscopy 30, 1453–1460 (2014).
49. Varma, H. S., Dadarya, B. & Vidyarthi, A. The new avenues in the management of osteo-arthritis of knee–stem cells. J. Indian Med. Assoc. 108, 583–585 (2010).

50. Vega, A. et al. Treatment of knee osteoarthritis with allogeneic bone marrow mesenchymal stem cells: a randomized controlled trial. Transplantation 99, 1681–1690 (2015).

51. Wakisani, S. et al. Human autologous culture expanded bone marrow mesenchymal cell transplantation for repair of cartilage defects in osteoarthritic knees. Osteoarthr. Cartil. 10, 199–206 (2002).

52. Wong, K. L. et al. Injectable cultured bone marrow-derived mesenchymal stem cells in varus knees with cartilage defects undergoing high tibial osteotomy: a prospective, randomized controlled clinical trial with 2 years’ follow-up. Arthroscopy 29, 2020–2028 (2013).

53. Biowal, S., Medhi, B. & Pandhi, P. Longterm efficacy of topical nonsteroidal anti-inflammatory drugs in knee osteoarthritis: metaanalysis of randomized placebo controlled clinical trials. J. Rheumatol. 33, 1841–1844 (2006).

54. Zhang, W. et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthr. Cartil. 18, 476–499 (2010).

55. Pham, T. et al. OMERACT-OARSI initiative: osteoarthritis research society international set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthr. Cartil. 12, 389–399 (2004).

56. Torres, L. et al. The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. Osteoarthr. Cartil. 14, 1033–1040 (2006).

57. Yamaguchi, S. et al. The effect of exercise on the early stages of mesenchymal stromal cell-induced cartilage repair in a rat osteochondral defect model. PLoS One 11, e0151580 (2016).

58. Yamaguchi, S. et al. Effect of low-intensity pulsed ultrasound after mesenchymal stromal cell injection to treat osteochondral defects: an in vivo study. Ultrasound Med. Biol. 42, 2903–2913 (2016).

59. Harada, Y. et al. Combination therapy with intra-articular injection of mesenchymal stem cells and articulated joint distraction for repair of a chronic osteochondral defect in the rabbit. J. Orthop. Res. 33, 1466–1473 (2015).

60. Rice, D. A., McNair, P. J., Lewis, G. N. & Dalbeth, N. Quadriceps arthrogenic muscle inhibition: the effects of experimental knee joint effusion on motor cortex excitability. Arthritis Res. Ther. 16, 502 (2014).

61. Freitag, J. et al. Mesenchymal stem cell therapy in the treatment of osteoarthritis: reparative pathways, safety and efficacy - a review. BMC Musculoskelet. Disord. 17, 230 (2016).

62. Tucker, J. D., Erickson, J. J., Goetz, L. L. & Elmore, L. W. Should clinical studies involving “regenerative injection therapy,” strive to incorporate a triad of outcome measures instead of only including clinical outcome measures? Osteoarthr. Cartil. 22, 715–717 (2014).

63. Chang, K. V. et al. Comparative effectiveness of platelet-rich plasma injections for treating knee joint cartilage degenerative pathology: a systematic review and meta-analysis. Arch. Phys. Med. Rehabil. 95, 562–575 (2014).

64. Tsujimoto, Y. et al. Protocol registration of systematic reviews published in high-impact factor journals: a meta-epidemiological study. J. Clin. Epidemiol. 84, 54–60 (2017).

65. McAlindon, T. E. et al. OARSI clinical trials recommendations: design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthr. Cartil. 23, 747–760 (2015).

66. Collins, N. J., Misra, D., Felson, D. T., Crosseley, K. M. & Roos, E. M. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). Arthritis Care Res. 63, S208–S228 (2011).

67. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Group, P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151, 264–269 (2009). W264.

68. Shamseer, L. et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 349, g7647 (2015).

69. Stroup, D. F. et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 283, 2008–2012 (2000).

70. Lo, G. H., LaValley, M., McAlindon, T. & Felson, D. T. Intra-articular hyaluronic acid in treatment of knee osteoarthritis: a meta-analysis. JAMA 290, 3115–3121 (2003).

71. Deeks, J. J. & Higgins, J. P. Statistical algorithms in review manager s. Statistical Methods Group of The Cochrane Collaboration. 1–11 (2010).

72. Anzures-Cabrera, J. & Higgins, J. P. Graphical displays for meta-analysis: an overview with suggestions for practice. Res. Synth. Methods 1, 66–80 (2010).

73. Cohen, J. A power primer. Psychol. Bull. 112, 155–159 (1992).

74. Egger, M., Davey Smith, G., Schneider, M. & Minder, C. Bias in meta-analysis detected by a simple, graphical test. BMJ 315, 629–634 (1997).

75. Duval, S. & Tweedie, R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–463 (2000).

76. Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. BMJ 327, 557–560 (2003).

© The Author(s) 2018