**Bond University**

**Research Repository**

# Effects of water-based exercise on bone health of middle-aged and older adults

## A systematic review and meta-analysis

Simas, Vini; Hing, Wayne; Pope, Rodney; Climstein, Mike

*Document Version:*
Publisher's PDF, also known as Version of record

[Link to publication in Bond University research repository.](#)

Open Access Full Text Article

REVIEW

# Effects of water-based exercise on bone health of middle-aged and older adults: a systematic review and meta-analysis

Vini Simas[1]
Wayne Hing[1]
Rodney Pope[1]
Mike Climstein[1,2]

[1]Water-Based Research Unit, Faculty of Health Sciences and Medicine, Bond Institute of Health and Sport, Bond University, Gold Coast, QLD, [2]Exercise, Health and Performance Research Group, Faculty of Health Sciences, The University of Sydney, Sydney, NSW, Australia

**Background:** Age-related bone loss is a major health concern. Only exercises associated with high-impact and mechanical loading have been linked to a positive effect on bone turnover; however, these types of exercises may not always be appropriate for middle-aged and older adults due to physical decline or chronic disorders such as osteoarthritis. Water-based exercise (WBE) has been shown to affect different components of physical fitness, has lower risks of traumatic fracture, and applies less stress to joints. However, the effects of WBE on bone health are unclear.

**Objective:** This study aimed to explore whether WBE is effective in preventing age-related bone deterioration in middle-aged and older adults.

**Methods:** A search of relevant databases and the references of identified studies was performed. Critical narrative synthesis and meta-analyses were conducted.

**Results:** Eleven studies, involving 629 participants, met all inclusion criteria. All participants were postmenopausal women. Eight studies compared WBE to a sedentary control group, and four studies had land-based exercise (LBE) participants as a comparison group. Meta-analyses revealed significant differences between WBE and control group in favor of WBE for changes in bone mineral density (BMD) at the lumbar spine (mean difference [MD] 0.03 g/cm$^2$; 95% confidence interval [CI]: 0.01 to 0.05) and femoral neck (MD 0.04 g/cm$^2$; 95% CI: 0.02 to 0.07). Significant differences were also revealed between WBE and LBE in favor of LBE for changes in lumbar spine BMD (MD −0.04 g/cm$^2$; 95% CI: −0.06 to −0.02). However, there was no significant difference between WBE and LBE for changes in femoral neck BMD (MD −0.03 g/cm$^2$; 95% CI: −0.08 to 0.01).

**Conclusion:** WBE may have benefits with respect to maintaining or improving bone health in postmenopausal women but less benefit when compared to LBE. Further research is required on this topic.

**Keywords:** aquatic exercise, bone mineral density, osteoporosis, preventive medicine, sports medicine

## Introduction

Age-related bone loss is a major health concern. Loss of bone mass and microarchitectural deterioration of bone tissue are directly related to a decrease in bone strength and subsequently increased fracture risk, which ultimately leads to conditions clinically known as osteopenia and osteoporosis.[1,2] Osteoporotic fractures have particular importance in public health and are considered one of the most common causes of disability, as well as a major contributor to medical care costs worldwide.[3] They are responsible for excess mortality, morbidity, chronic pain, reduction in quality of life, and admission to long-term care, significantly contributing to health and social care costs.[4] In Australia, it is estimated that osteopenia and osteoporosis affect ~7.5 million

Correspondence: Vini Simas
Water-Based Research Unit, Bond Institute of Health and Sport, Faculty of Health Sciences and Medicine, Bond University, Gold Coast, 2 Promethean Way, Robina, Gold Coast, QLD 4226, Australia
Tel +61 405 617 133
Email vpsimas@gmail.com

people, with one fragility fracture occurring every 3.6 minutes, which amounts to ~400 per day.[5–7] The estimated total number of osteoporotic new fractures and refractures over the period 2012–2022 is predicted to be in excess of 1.6 million, with an estimated total direct and indirect cost to the Australian government, community, and individuals of AU$33.6 billion in this period.[8] Over this period, it is also projected that ~150,000 fractures could be prevented, with an annual saving ranging from AU$140 million to AU$187 million.[8] The residual lifetime risk of osteoporotic fractures for women aged 50 years is estimated to be >40% and represent 80% of all fractures in the population over this age.[9] For men aged 60 years, the residual lifetime fracture risk is estimated to be ~30%.[10]

The most common sites of osteoporotic fractures are the hips, spine, and wrists. Hip fractures account for the majority of direct medical costs and are also an important contributor to long-term disability, with ~30% of older adults with a history of hip fracture not reaching their prefracture level of functioning 1 year following a fracture.[11] In addition to this, in the year following a hip fracture, there is a twofold increase in mortality,[12] estimated to be ~30%, and it is higher among male patients (37.5%).[13] Vertebral osteoporotic fractures are often asymptomatic, therefore escaping clinical diagnosis; however, when compared to other types of fragility fractures, they are associated with higher comorbidity, higher incidence of hospitalization, and longer hospital stays.[14] In addition, they have been strongly related to subsequent fractures and mortality.[14,15] The residual lifetime risk of vertebral osteoporotic fractures is 8.6% for men aged ≥45 years and 15.4% for women.[14] Distal radius fractures (occurring at the wrist) are more prevalent in women aged 45–65 years, and the most common mechanism of these fractures is direct trauma.[16] Although fractures of the distal radius are considered to cause the least morbidity of all fragility fractures, these fractures are regarded as an important predictor of subsequent fractures and mortality.[17]

Even though the majority (60–80%) of the variation in bone strength is attributed to genetics,[18–21] bone is considered a dynamic tissue, exhibiting continuous remodeling activity. This remodeling process is mediated by osteoblasts, which are cells responsible for bone formation, and osteoclasts, which are cells responsible for bone resorption, causing bone loss. The remodeling process is capable of adapting and responding to various stimuli.[22–24] On this basis, it is estimated that lifestyle and environmental factors, such as nutrition, alcohol intake, smoking, and skeletal loading, contribute to 20–40% of the variation in bone quality.[25] It is well known

that prolonged periods of inactivity and unloading of the skeleton have a negative effect on bone mass, accelerating bone loss.[26] In addition, lean body mass and skeletal muscle mass are strongly related to bone mineral density (BMD).[27–29] It is also well documented that muscle contractions can increase loads on bones, generating stress and strain reactions in bone tissue,[30–32] and that dynamic loading has a more positive effect on bone tissue than static loading.[33]

Many efforts have been made to investigate nonpharmacological approaches for achieving an osteogenic (bone-producing) effect. It is well known that avoidance of tobacco and adequate serum levels of calcium and vitamin D are essential for bone health.[34–36] Physical activity has been shown to be an effective nonpharmacological approach to improve bone mass; however, not all types of exercise have been definitively shown to promote positive effects on bone metabolism.[37] In research to date, only impact weight bearing and high-impact progressive resistance training activities have a strong level of evidence indicating a positive osteogenic effect.[38–43] However, it is well known that aging can also be associated with physical decline, including conditions such as joint limitations and chronic pain, and, therefore, high-impact exercise is not always indicated or appropriate for middle-aged and older adults.

Exercise executed in the water environment, often referred to as water-based exercise (WBE), presents lower risks of traumatic fracture, and the joints are exposed to less stress and impact (via reduced loading due to buoyancy), when compared to land-based exercise (LBE), such as running, resistance training, and strength training. Besides this, WBE has been highly recommended for older people, especially those with disability, due to the reduced pain[44] and increased security it can provide,[45] in addition to providing additional benefits for neuromuscular and functional fitness,[45] and cardiometabolic health.[46] Furthermore, considering the potential for a reduction in the prevalence of pain and injuries, the dropout rate among subjects participating in WBE may be lower than that for some land-based activities. Finally, some older adults may simply enjoy WBE or wish to participate due to social reasons. In WBE, increased muscular demands are often necessary in order to overcome water resistance. For instance, Chevutschi et al[47] demonstrated that walking in water at an umbilical level increased the activity of the erector spinae and activated the rectus femoris to levels near to or higher than walking on dry ground. Therefore, considering the muscle demands and the dynamic component of WBE, there might be adequate stimulus to generate osteogenic stress and strain reactions in bones.

However, the literature is inconsistent in its reports of the effects of WBE on bone health of middle-aged and older adults. Some observational studies that have investigated swimmers have reported that participants have similar, or sometimes lower, BMD when compared to sedentary controls, indicating that swimming is associated with a similar or greater risk of bone deterioration and its consequences when compared to a sedentary lifestyle.[48–50] Velez et al[50] reported that mature-aged males who restricted their physical activity to only swimming had a 10% higher prevalence of osteoporosis when compared to sedentary age- and sex-matched controls. Conversely, in a cross-sectional analysis, Balsamo et al[51] concluded that aquatic exercise might be an effective nonpharmacological strategy to prevent bone loss in postmenopausal women. In addition to this, Gomez-Bruton et al[52] conducted a systematic review analyzing the effects of swimming on bone tissue, analyzing 64 studies assessing children, adolescents, adults, and elderly populations. It was reported that swimming had no negative influence on bone tissue and might have benefits on bone health later in life.

To date, a consensus regarding the effects of exercise practiced in water on bone health has not been reached, and a comprehensive literature search conducted by the authors identified no systematic review of the effects of WBE other than swimming. Therefore, the effects of exercise undertaken in a water environment on bone health of middle-aged and older adults remain uncertain.

This systematic review and meta-analysis aimed to answer the following question: is WBE effective in preventing age-related bone deterioration in middle-aged and older adults? The objective of the review was to assess the effect of WBE interventions in preventing age-related bone deterioration when compared to a sedentary lifestyle or other forms of exercise.

Findings of this systematic review and meta-analysis are expected to contribute to the knowledge of health-care professionals involved in this field with regard to the effectiveness of WBE, so that they can provide alternative recommendations regarding exercise types that can maintain or even enhance bone health and reduce the risk of fracture among their patients or clientele.

## Methods

The review was conducted as a systematic review of relevant studies, incorporating both a critical narrative synthesis and a meta-analysis. The design of this study was guided by consideration of the Cochrane Handbook for Systematic Reviews of Interventions[53] and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.[54] The methods and eligibility criteria for included studies were detailed in advance in a protocol registered at the international database of prospectively registered systematic reviews in health and social care, PROSPERO[55] (registration number: CRD42015026685).

## Eligibility criteria

To be included in the review, studies were required to be full-length research articles, published in academic journals or online (e-publication ahead of print), and no limits were set on language or date of publication. Only clinical trials (randomized [RCTs] or nonrandomized controlled trials) and prospective observational studies were considered for inclusion, with no limits set on length of follow-up. Studies were also considered only if they analyzed human subjects, either male or female, and if participants were aged 45 years or older, asymptomatic, and free living in the community. Participants in eligible studies could be healthy individuals or individuals with diagnosed osteopenia or osteoporosis; however, studies involving participants with other known health disorders or restrictions on participation in physical activities were ineligible. In addition, studies included in the systematic review were required to have a type of WBE or physical activity as the only intervention or exposure in at least one group and a comparison group, such as people undertaking other types of exercise or sedentary controls. If any medication or supplements were given to one group, the study was only considered if the medication or supplement was also given to all other groups, using the same dosage. Eligible studies assessed BMD by dual-energy X-ray absorptiometry (DEXA). The primary outcomes of interest in this review were BMD and bone mineral content (BMC) measured by DEXA, measured in at least one clinical site (lumbar spine [LS], proximal femur, total hip, or wrist); bone metabolism measured by serum biomarkers; incidence rates of bone fractures; minor adverse events, including falls; and serious adverse events, including death. The secondary outcomes of interest were muscle strength, flexibility, balance, and compliance with the intervention.

The following exclusion criteria were applied during study selection: publication types other than full-length journal articles, such as letters to the editor, conference abstracts, conference papers or book chapters; unpublished papers; studies using a descriptive or retrospective design; studies that did not evaluate human subjects; studies that did not evaluate middle-aged or older adults or that evaluated middle-aged or older adults together with other age groups

without reporting the results separately; studies involving participants with medical disorders other than osteopenia or osteoporosis; studies that did not have WBE as the sole intervention in at least one group; studies that did not have a comparison group; studies that did not have BMD as an outcome; and studies that did not measure BMD by DEXA.

## Search methods

To identify relevant studies, a multistep search was conducted in October 2015, without any limits on publication date, in the following databases: PubMed/MEDLINE, the Cochrane Library, EMBASE, SPORTDiscus, CINAHL, ScienceDirect, Scopus, AUSPORT, and PEDro. In addition, hand searches of reference lists of included articles were also performed to identify additional studies and data that met criteria for inclusion. The search strategy was kept as broad as possible, with identification of articles achieved by the use of specific text words, without using truncation, wildcards, or any other limits. Search strategies for all databases were tailored to the nuances of the respective database and are available upon request.

## Data collection and analysis

Search results were imported into reference management software (EndNote),[56] where duplicate records were removed. Titles and abstracts were then screened, in order to exclude studies that were clearly ineligible. After initial screening, potentially eligible studies were retrieved for full-text eligibility assessment. The selection process applied to the full-text study reports was based upon the eligibility criteria discussed earlier, including types of interventions, types of outcome measures, types of participants, and types of studies. Disagreements regarding assessed eligibility were resolved by consensus and reasons for exclusion of studies were documented. The results of the entire search, screening, and selection process were recorded in a PRISMA diagram (Figure 1).[54]

Data were extracted and tabulated from all included papers using a standardized data extraction tool (The Cochrane Consumers and Communication Review Group).[53] Data extracted from each paper included specific details of title, authors, source, year of publication, study design, participants, the intervention, the comparison groups, length of follow-up, and data related to the primary and secondary outcomes of interest for this review.

Risk of bias was assessed for each included study using the Cochrane Collaboration's Risk of Bias tool.[57] The following elements that potentially affect risk of bias were addressed: random sequence generation (selection bias),

allocation concealment (selection bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and other sources of bias (comparability of treatment and control group [CG] at entry, appropriateness of duration of follow-up). The risk of bias in the included studies was narratively described, and then each item was assigned a judgment: "low", "high," or "unclear" risk of bias. Nonrandomized controlled trials (quasi-experimental studies) and prospective observational studies were assessed and reported as being at a high risk of bias on the random sequence generation and allocation concealment items of the risk of bias tool.

Quantitative data were analyzed using the Cochrane software Review Manager (RevMan, version 5.3),[58] where outcomes were reported in at least two studies. Effect sizes for continuous outcomes were calculated as mean differences (MDs) or, if different scales had been used, as standardized mean differences (SMD), each with 95% confidence intervals (CIs), using a random-effects model. Missing data and attrition rates were assessed for each of the included studies and were reported as the proportion of commencing participants included in the final analysis. Intention-to-treat analysis of reported data from each included study was applied when extracting data for the meta-analysis. That is, each participant was included in the group to which they were randomized, and all randomized participants were included in the analysis. Heterogeneity was assessed using the standard $\chi^2$ test and $I^2$ value.[53] Heterogeneity was considered statistically significant at $P < 0.10$. $I^2$ values between 0% and 30% were considered minimal, 30%–50% moderate, 50%–90% substantial, and >90% considerable. The overall treatment or intervention effect was calculated for each outcome measure in each included study. The effect of treatment or intervention on each outcome measure was calculated as the difference between the intervention and CGs in the change in measured outcome from baseline to the end of follow-up. For each outcome measure, variance was estimated based on the standard deviation (SD) of the MD between baseline and follow-up. When this value was not available and was not supplied by the respective study authors following a written request, we used the SD calculated from the $P$-value for the differences between mean values in the groups.[53] When the $P$-value was not available, we imputed the highest SD available from other studies included in the review.

## Results

### Search, screening, and selection results

The search of electronic databases retrieved 12,271 records, with an additional 25 articles identified by

**Identification**

Records identified through database searching
(n = 12,617)

• PubMed/MEDLINE = 2,666
• Cochrane = 447
• Embase = 2,647
• SPORTDiscus = 676
• CINAHL = 539
• Scopus = 2,714
• AUSPORT = 7
• ScienceDirect = 816
• PEDro = 2,105

Records identified through other sources
(n = 25)

**Screening**

Duplicates removed
(n = 4,819)

Records screened
(n = 7,823)

Records excluded
(n = 7,737)

• Not in humans = 630
• Did not analyze water-based exercise and bone health = 6,576
• Did not analyze middle-aged or older adults = 226
• Population with medical disorders = 279
• Other types of publications = 26

**Eligibility**

Full-text articles assessed for eligibility
(n = 86)

Full-text articles excluded, with reasons
(n = 75)

• Did not analyze water-based exercise and bone health = 8
• Did not analyze middle-aged or older adults = 31
• Did not analyze bone health via DEXA = 2
• Did not have a comparison group = 1
• Other types of study design:
 – Narrative review = 16
 – Systematic review = 1
 – Retrospective = 3
 – Cross-sectional = 13

**Included**

Studies included in critical narrative synthesis and meta-analysis
(n = 11)

**Figure 1** PRISMA flow diagram.
**Abbreviation:** PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analysis.

searching references of potentially eligible articles. After removing duplicates, 7,823 articles remained to be screened by title and abstract, with 7,737 of these being excluded because they clearly did not meet eligibility criteria and 86 articles then remaining to be assessed for eligibility in full text. From these full-text articles, 11 articles[59–69] that met eligibility criteria were identified and included in this review. Results from the search, screening, and selection processes are summarized in a

PRISMA flow diagram (Figure 1).[54] Table 1 summarizes the characteristics of the 11 included studies.

## Description of included studies

Of the 11 included studies, five were RCTs and six quasi-experiments (QEs). A total of 629 participants were divided into three groups: participants who performed WBE (n=344), participants who performed LBE (n=82), such as resistance training and strength training, and a sedentary

**Table I** General characteristics of included studies

| Study ID | Design | Duration, months | Number of participants | Losses | Participants | Age, years |
|---|---|---|---|---|---|---|
| Borba-Pinheiro et al[59] | QE | 12 | 35 | NR | • Sex: females<br>• Characteristics: women with osteoporosis/osteopenia, being treated with alendronate, no previous history of fractures and no history for at least 1 year of regular practice of PA, in good physical and mental health | 45.6–64.5 |
| Borba-Pinheiro et al[60] | QE | 12 | 84 | NR | • Sex: females<br>• Characteristics: volunteers with osteoporosis and/or osteopenia in at least one of the DEXA measures, undergoing treatment with alendronate sodium (70 mg/week) and/or vitamin D, and no history of fractures, in good physical and mental health | 49–61.8 |
| Kemper et al[61] | RCT | 6 | 30 | 7 (2 SWM/5 RTG) | • Sex: females<br>• Characteristics: sedentary postmenopausal women | 63.9 ± 6.49 (mean ± SD) |
| Moreira et al[62] | RCT | 6 | 108 | 8 (5 AEG/3 CG) | • Sex: females<br>• Characteristics: physical activity status classified as sedentary by the IPAQ short version questionnaire, postmenopausal for at least 5 years and cognitive function allowed to understand and respond to the authors' questions and commands during the questionnaires, no physical conditions that might affect performance during aquatic exercises; no chronic kidney disease; no history of recent hip fracture (in the last 2 years); no dependency on alcohol or illicit drugs; no chronic therapy with corticosteroids, bisphosphonates, calcitonin, calcium, vitamin D, and its metabolites; no use of estrogen, selective estrogen receptor modulators, and strontium in the earlier 6 months; no use of any medications that might interfere with vitamin D metabolism; systolic blood pressure <200 mmHg and/or diastolic blood pressure <100 mmHg | 58.8 ± 6.4 (mean ± SD) |
| Murtezani et al[63] | RCT | 10 | 64 | 3 (2 LBE/1 AE) | • Sex: females<br>• Characteristics: women recently diagnosed (within the past 6 months) with osteoporosis on account of a DEXA scan T score below −2.5, aged 50–70 years, who had no history of vertebral fractures or lower extremity fractures, did not have endoprostheses or fixation materials and were capable of signing written informed consent | 50–70 |

| Water-based exercise | Comparison groups | BMD measurement | Secondary outcomes | Medications/ supplements | Adverse events | Compliance/ adherence |
|---|---|---|---|---|---|---|
| • Exercise: hydrogymnastics<br>• Frequency: 3 times/week, 60 min, six bimonthly cycle<br>• Intensity: Borg scale, 12 during the first 2 months and 14–16 during the rest of the study<br>• Setting: 12 m section of a 25 m pool | • RTG<br>• JUG<br>• CG | • Equipment: DEXA Lunar DPX<br>• Variables: BMD (g/cm$^2$)<br>• Regions: lumbar spine (L2–L4), neck of the femur, greater trochanter and Ward's triangle | • Body balance<br>• Quality of life | Alendronate sodium 70 mg/week | NR | NR |
| • Exercises: hydrogymnastics<br>• Frequency: three 60-min sessions per week on alternate days<br>• Intensity: Borg scale, 11–12 during the first 2 months and 13–16 during the rest of the study<br>• Setting: 12 m section of a 25 m pool | • CG | • Equipment: DEXA Lunar DPX<br>• Variables: BMD (g/cm$^2$)<br>• Regions: lumbar spine (L2–L4), femoral neck and greater trochanter | • Quality of life | Alendronate sodium 70 mg/week and/or vitamin D3 5600 IU/week | NR | NR |
| • Exercises: swimming<br>• Frequency: 3 days/week, 1 hour per session<br>• Intensity: sessions began with moderate intensity activities (60% HRR) and reached high intensity activities (90% HRR)<br>• Setting: pool with 1.50 m depth, and the water temperature during sessions remained between 27°C and 29°C | • RTG | • Equipment: DEXA Lunar DPX-IQ<br>• Variables: BMD (g/cm$^2$)<br>• Regions: lumbar spine (L2–L4) and femoral neck | • Body composition | – | NR | >75% of sessions |
| • Exercises: hydrogymnastics<br>• Frequency: 3 times/week, sessions lasted from 50 to 60 min<br>• Intensity: level 6 in Borg scale (~60% of MHR) during 16 min of the session in weeks 5–9, level 7 in Borg scale (~70 % of MHR) during 13 min of the session in weeks 10–14, level 8 in Borg scale (~80 % of MHR) during 9 min of the session in weeks 15–19, and level 9 in Borg scale (~90 % of MHR), during 7 min of the session in weeks 20–24<br>• Setting: covered swimming pool, with depth varying between 1.10 and 1.30 m and water temperature between 30°C and 31°C | • CG | • Equipment: DEXA Hologic QDR<br>• Variables: BMD (g/cm$^2$)<br>• Regions: lumbar spine (L1–L4), femoral neck, total femur, total body | • Biomarkers of bone turnover (iPTH, P1NP, and CTx) | Daily supplement of 500 mg of elementary calcium and 1,000 IU of vitamin D (cholecalciferol), combined in the same pill | NR | 92.6% (95% CI, 85–98%) |
| • Exercises: hydrogymnastics<br>• Frequency: 3 times/week, 35 min<br>• Intensity: participants were instructed to exercise at an intensity that was moderate to hard (12–14 on the 20-point Borg perceived exertion scale)<br>• Setting: not stated (only reported that water temperature was 30°C) | • LBE | • Equipment: DEXA GE Lunar Prodigy<br>• Variables: BMD (g/cm$^2$)<br>• Regions: lumbar spine (L2–L4) | • Muscle strength<br>• Flexibility<br>• Balance<br>• Gait time<br>• Pain | Ca (1,000 mg daily) and vitamin D (800–1,000 IU daily) | NR | NR |

(*Continued*)

**Table 1** (Continued)

| Study ID | Design | Duration, months | Number of participants | Losses | Participants | Age, years |
|---|---|---|---|---|---|---|
| Novaes et al[64] | QE | 6 | 31 | NR | • Sex: females<br>• Characteristics: living independently in the community, aged ≥55 years, postmenopausal status, being without contraindications to physical activity, and not reporting history of regular structured exercise | 66.9 ± 6.1 (mean ± SD) |
| Pernambuco et al[65] | RCT | 8 | 84 | 17 (6 AEG/11 CG) | Sex: females<br>Characteristics: aged ≥60 years, with low BMD and no neurological disorders who did not exercise regularly for a minimum of 6 months prior to the study and did not suffer from metabolic or endocrine disorders | 60–77 |
| Rotstein et al[66] | QE | 7 | 35 | 5 (all from intervention group) | Sex: females<br>Characteristics: nonsmokers did not suffer from thyroid gland problems or hypertension and did not suffer from osteoporosis (baseline bone density was >55% of the mean bone density for the normal population of the subject's age), not taking any of the following medications: B complex, Betaxolol ophthalmic suspension, anastrozole, brotizolam, glucosamine, chondroitin, ofloxacin, atenolol, cilazapril, estradiol, norethisterone, pravastatin, losartan, amlodipine, and aspirin | 50–65 |
| Tsukahara et al[67] | QE | 12 | 97 | 32 (7 from veterans group and 25 from newcomers group) | • Sex: females<br>• Characteristics: healthy postmenopausal women | 59.75–65.08 |
| Vanaky et al[68] | RCT | 3 | 20 | NR | • Sex: females<br>• Characteristics: nonsmoker females aged between 50 and 70 years; postmenopausal for at least 12 months; not institutionalized; and having no contraindication to undertake physical exercises without close medical supervision, hormone therapy, and calcium consumption and without cardiovascular and thyroid history | 50–70 |
| Wu et al[69] | QE | 24 | 41 | | • Sex: females<br>• Characteristics: postmenopausal | SWM – 59.5 ± 6.1<br>CG – 59.3 ± 5.2 |

**Abbreviations:** BMD, bone mineral density; QE, quasiexperiment; NR, not reported; RTG, resistance training group; JUG, judo group; CG, control group; DEXA, dual-energy X-ray absorptiometry; SWM, swimming group; HRR, heart rate reserve; SD, standard deviation; P1NP, procollagen type 1 amino-terminal propeptide; CTx, carboxy-terminal cross-linking telopeptide of type 1 collagen; AEG, aquatic exercises group; RCT, randomized controlled trial; BMC, bone mineral content; LBE, land-based exercise; ST, strength training ; CI, confidence interval; min, minutes; PA, physical activity; IPAQ, international physical activity questionnaire; AE, aquatic exercise; iPTH, intact parathyroid hormone; MHR, maximum heart rate.

| Water-based exercise | Comparison groups | BMD measurement | Secondary outcomes | Medications/ supplements | Adverse events | Compliance/ adherence |
|---|---|---|---|---|---|---|
| • Exercises: hydrogymnastics<br>• Frequency: 3 times/week, sessions lasting 45 min<br>• Intensity: participants were instructed to attain 70%–80% of their work heart rate<br>• Setting: public indoor swimming pool with a water depth of 1.20–1.40 m and a water temperature of 30.5°C | • ST | • Equipment: DEXA Hologic<br>• Variables: BMD (g/cm²)<br>• Regions: left femoral neck and lumbar spine (L1–L4) | – | – | – | >85% |
| • Exercises: hydrogymnastics<br>• Frequency: twice weekly, 50 min<br>• Intensity: not reported<br>• Setting: swimming pool 25 m long and 1.40 m deep | • CG | • Equipment: DEXA Lunar DPX-L<br>• Variables: BMD (g/cm²)<br>• Regions: lumbar spine (L2–L4) and right total femur | • Bone formation (serum osteocalcin)<br>• Functional autonomy | Volunteers with osteoporosis took alendronate sodium (70 mg) once a week and vitamin D3 once a day, while those with osteopenia used only vitamin D3 | NR | NR |
| • Exercise: hydrogymnastics<br>• Frequency: three 1 hour sessions per week<br>• Intensity: 12–16 on Borg scale<br>• Setting: pool, water temperature was 32°C, and all activities were conducted with the water at chest level | • CG | • Equipment: DEXA Lunar<br>• Variables: BMD (g/cm²) and BMC (g)<br>• Regions: lumbar spine (L2–L4) and femoral neck | – | – | NR | NR |
| • Exercise: hydrogymnastics and swimming<br>• Frequency: once/week, 45 min/ session<br>• Intensity: level of activity had two maximum working heart rate peaks (~120 beats/min)<br>• Setting: sports club, warm water (28–29°C) | • Newcomers<br>• CG | • Equipment: DEXA Hologic QDR<br>• Variables: BMD (g/cm²)<br>• Regions: lumbar spine (L1–L4) | – | – | NR | NR |
| • Exercise: hydrogymnastics<br>• Frequency: 3 times/week, 1 h, 15 min/ session (after the second week)<br>• Intensity: the intensity of the jumps was adjusted to ~60% of the HRR. During the second week, the intensity was increased to 80% of the HRR<br>• Setting: pool | • CG | • Equipment: DEXA<br>• Variables: BMD<br>• Regions: lumbar spine (L2–L4) and femoral neck | • Functional fitness | – | NR | NR |
| • Exercise: swimming<br>• Frequency: 1.5 times/week, 1 h/ session<br>• Intensity: NR<br>• Setting: pool | • CG | • Equipment: DEXA Norland<br>• Variables: BMD (g/cm²)<br>• Regions: lumbar spine (L1–L4) and proximal femur | • Leg extension power | – | – | – |

CG (n=203). All participants in the studies were postmeno-pausal women. Four studies reported that the participants were previously sedentary.[59,61,62,64] Regarding bone health, four studies recruited participants with low BMD (osteope-nia or osteoporosis)[59,60,63,65] and one recruited women with normal BMD.[66] Groups from one study received alendronate sodium,[59] groups from two studies received a combination of alendronate sodium and vitamin D,[60,65] and groups from another two studies received a combination of vitamin D and calcium.[62,63] The studies were conducted in Brazil (n=5),[59–62,65] Japan (n=2),[67,69] Kosovo (n=1),[63] Israel (n=1),[66] Iran (n=1),[68] and Portugal (n=1).[64] Nine studies were published in Eng-lish,[59,60,62–68] one article was translated from Portuguese,[61] and one from Japanese.[69] The length of the exercise interventions varied in the included studies: one study conducted the inter-vention for 24 months,[69] three for 12 months,[59,60,67] one for 10 months,[63] one for 8 months,[65] one for 7 months,[66] three for 6 months,[61,62,64] and one for 3 months.[68] The frequency and duration of the sessions also varied in the included studies, ranging from once a week to three times a week, and each session lasted from 35 to 75 minutes. The content of the train-ing sessions for WBE groups comprised hydrogymnastics in eight studies[59,60,62–66,68] and swimming in two studies.[61,69] One study combined both hydrogymnastics and swimming during the sessions.[67] Nine studies reported that exercise intensity was moderate to vigorous,[59–64,66–68] with the level of intensity determined by either heart rate or Borg scale. Four studies involved LBE groups as comparison groups, and the LBE training sessions consisted of resistance training,[59,61] strength training,[64] a mixture of aerobics and resistance training,[63] and judo.[59] Eight studies compared WBE to a sedentary CG.[59,60,62,65–69] One study included both WBE and LBE, as well as a CG.[59]

## Risk of bias in included studies

The judgment about each risk of bias item for each included study is presented in Figure 2, and the percentages of all included studies deemed to be at low risk, unclear risk, or high risk of bias based on each bias item are depicted in Figure 3.

### Random sequence generation and allocation concealment (selection bias)

All nonrandomized studies (QE) were classified as "high risk" for both "random sequence generation" and "allocation concealment" items. Of the five RCTs included in the review, three reported adequate sequence generation and were classi-fied as being at "low risk" of bias on this item.[62,63,65] The other two studies[61,68] reported that participants were randomized into groups; however, methods of randomization were not
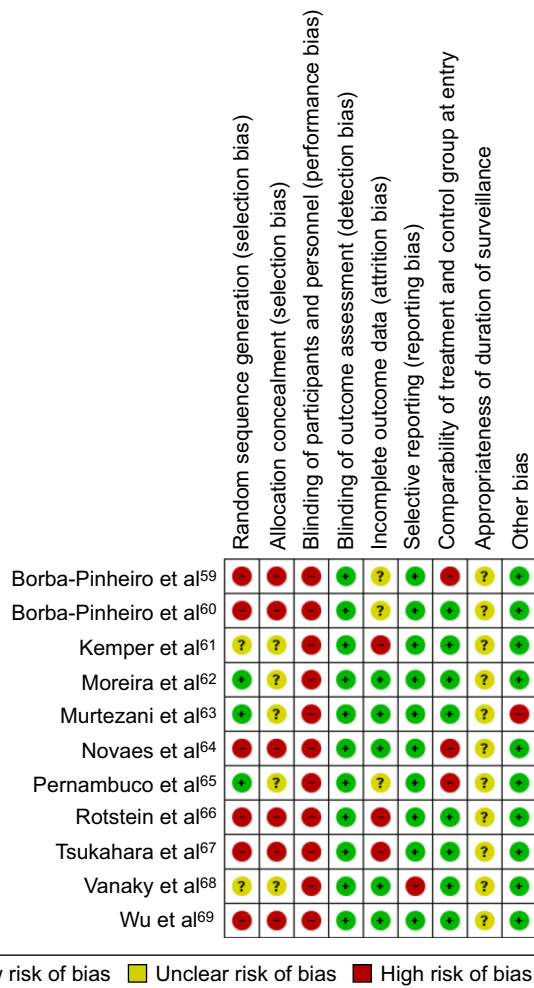


**Figure 2** Risk of bias summary, by item and study.

described and they were classified as being at "unclear risk" of bias on this item. None of the included RCTs described allocation concealment and so all were classified as being at "unclear risk" of bias for this item.

### Blinding of participants and personnel (performance bias)

All studies were classified as being at "high risk" for perfor-mance bias, as none of the studies reported any attempt to blind participants and personnel (such as exercise instructors and researchers) to group allocations.

### Blinding of outcome assessment (detection bias)

Considering the objective nature of the primary outcomes of interest, all studies were judged to be at "low risk" for detection bias.

### Incomplete outcome data (attrition bias)

Five studies were considered to be at "low risk" of attri-tion bias as they either reported data for all participants or

Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)
Comparability of treatment and control group at entry
Appropriateness of duration of surveillance
Other bias

0%    25%    50%    75%    100%

Low risk of bias    Unclear risk of bias    High risk of bias

**Figure 3** Risk of bias graph, by item.

appropriately addressed incomplete outcome data.[62–64,68,69] Three studies were judged to be at "unclear risk"[59,60,65] and three at "high risk".[61,66,67] The study conducted by Kemper et al[61] reported over 30% attrition for the LBE group and ~13% for the WBE group, and those lost to follow-up were not accounted for in the final analysis. Rotstein et al[66] reported 20% attrition in the WBE group, with no reasons mentioned and again the analysis did not account for those lost to follow-up. In the study conducted by Tsukahara et al,[67] there was an attrition rate of over 62% in the WBE group, with no reasons mentioned and no adjustment of the analysis to account for the losses.

### Selective reporting (reporting bias)

In all but one study, the primary outcome was reasonably well reported. Vanaky et al[68] reported their findings in a table that was poorly formatted and one of the reported results made no sense, and, therefore, this study was classified as presenting a "high risk" of reporting bias.

### Comparability of groups at entry

Three studies[59,64,65] were judged to be at "high risk" of bias due to inadequate group comparability at entry. All other studies were judged to be at "low risk" of bias on this item.

### Appropriateness of duration of follow-up

All studies were classified as being at "unclear risk" of bias stemming from lack of adequate duration of follow-up, as they only reported immediate postintervention data.

### Other bias

In the study conducted by Murtezani et al,[63] the LBE group engaged in longer and more frequent training sessions than the WBE group. In the discussion section of that paper, it was mentioned that the WBE group exercised twice a week for 30 minutes, whereas the LBE group exercised thrice a week for 55 minutes. Therefore, this study was judged to be at "high risk" of bias due to the different doses of exercises provided to the groups. All other studies appeared to be free from other obvious sources of bias.

## Primary outcomes
### BMD

All studies reported BMD for at least one clinical site. All studies reported BMD for the LS, eight reported BMD for the femoral neck (FN),[59–62,64,66,68,69] four reported BMD for the greater trochanter (GT),[59,60,62,69] two reported BMD for Ward's triangle (WT),[59,69] and two reported BMD for the total femur (TF).[62,65]

### LS BMD

LS BMD increased in participants performing WBE in 10 studies; however, this change was statistically significant in only one study.[68] Wu et al[69] reported a nonsignificant decrease in LS BMD in the WBE group. All eight studies that included a CG reported a nonsignificant decrease in LS BMD for this group.[59,60,62,65–69] Of the four studies reporting an LBE group, three reported a statistically significant increase in LS BMD in this group.[59,63,64] Kemper et al[61] reported a nonsignificant decrease. When comparing the results between groups, eight studies compared WBE and CG, and two described a statistically significant difference in change in LS BMD, in favor of the WBE group.[66,68] In the comparison between WBE and LBE, two studies described a statistically significant difference between these exercise types in effects on LS BMD, in favor of the LBE.[63,64] The results of a meta-analysis

| Study or subgroup | WBE Mean | SD | Total | CG Mean | SD | Total | Weight | Mean difference IV, random, 95% CI | Mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **1.1.1 24 months** | | | | | | | | | |
| Wu et al[69] | −0.018 | 0.347 | 22 | −0.016 | 0.347 | 19 | 1.2% | −0.00 [−0.22, 0.21] | |
| **Subtotal (95% CI)** | | | **22** | | | **19** | **1.2%** | **−0.00 [−0.22, 0.21]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 0.02 (P = 0.99) | | | | | | | | | |
| **1.1.2 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | 0.034 | 0.087 | 8 | −0.024 | 0.087 | 7 | 7.0% | 0.06 [−0.03, 0.15] | |
| Borba-Pinheiro et al[60] | 0.034 | 0.135 | 43 | −0.024 | 0.135 | 41 | 16.3% | 0.06 [−0.00, 0.12] | |
| Tsukahara et al[67] | 0.004 | 0.347 | 40 | −0.021 | 0.347 | 30 | 2.0% | 0.03 [−0.14, 0.19] | |
| **Subtotal (95% CI)** | | | **91** | | | **78** | **25.3%** | **0.06 [0.01, 0.10]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 0.14, df = 2 (P = 0.93); $I^2$ = 0% | | | | | | | | | |
| Test for overall effect: Z = 2.34 (P = 0.02) | | | | | | | | | |
| **1.1.3 8 months** | | | | | | | | | |
| Pernambuco et al[65] | 0.026 | 0.347 | 42 | −0.003 | 0.347 | 42 | 2.5% | 0.03 [−0.12, 0.18] | |
| **Subtotal (95% CI)** | | | **42** | | | **42** | **2.5%** | **0.03 [−0.12, 0.18]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 0.38 (P = 0.70) | | | | | | | | | |
| **1.1.4 7 months** | | | | | | | | | |
| Rotstein et al[66] | 0.01 | 0.039 | 25 | −0.03 | 0.039 | 10 | 66.5% | 0.04 [0.01, 0.07] | |
| **Subtotal (95% CI)** | | | **25** | | | **10** | **66.5%** | **0.04 [0.01, 0.07]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 2.74 (P = 0.006) | | | | | | | | | |
| **1.1.5 6 months** | | | | | | | | | |
| Moreira et al[62] | 0.002 | 0.286 | 64 | −0.012 | 0.286 | 44 | 4.5% | 0.01 [−0.10, 0.12] | |
| **Subtotal (95% CI)** | | | **64** | | | **44** | **4.5%** | **0.01 [−0.10, 0.12]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 0.25 (P = 0.80) | | | | | | | | | |
| **Total (95% CI)** | | | **244** | | | **193** | **100.0%** | **0.04 [0.02, 0.07]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 0.92, df = 6 (P = 0.99); $I^2$ = 0% | | | | | | | | | |
| Test for overall effect: Z = 3.52 (P = 0.0004) | | | | | | | | | |
| Test for subgroup differences: $\chi^2$ = 0.78, df = 4 (P = 0.94), $I^2$ = 0% | | | | | | | | | |

**Figure 4** Forest plot of comparison between WBE and CG for changes in LS BMD (mean difference in g/cm$^2$).
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; LS, lumbar spine; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.

comparing the effects of WBE and CG on LS BMD are shown in Figure 4. The results revealed a significant difference between the groups in favor of WBE (MD 0.04 g/cm$^2$; 95% CI 0.02 to 0.07; P = 0.0004; $I^2$ = 0%). In this meta-analysis, we excluded the study conducted by Vanaky et al,[68] due to its high risk of reporting bias, but a subsequent sensitivity analysis indicated that its inclusion in the analysis would not have affected the overall result anyway. For the comparison of the effects of WBE and LBE interventions on LS BMD, results revealed a significant difference between the interventions in favor of LBE (MD −0.04 g/cm$^2$; 95% CI −0.06 to −0.02; P < 0.00001; $I^2$ = 0%), as shown in Figure 5.

## FN BMD

Of the eight studies that examined FN BMD, five reported an increase in this value for the WBE group;[61,62,66,68,69] however, only two studies reported a statistically significant change.[68,69] Two studies described a nonsignificant decrease in FN BMD

in the WBE group,[59,60] and one study reported the same value at baseline and postintervention time points.[64] All six studies that assessed a CG reported a nonsignificant decrease in FN BMD in this sedentary group.[59,60,62,66,68,69] Of the three studies that assessed FN BMD in the LBE group, two studies described an increase,[59,64] which was statistically significant in one study,[64] and one study described a nonsignificant decrease.[61] When WBE was compared to CG, two studies reported statistically significant differences in FN BMD changes, in favor of WBE.[68,69] In the comparison between WBE and LBE, two studies reported statistically significant differences in FN BMD changes, in favor of LBE.[59,64] Figure 6 details the results of the meta-analysis comparing FN BMD changes in WBE and CG, showing that there was a statistically significant difference in favor of WBE (MD 0.03 g/cm$^2$; 95% CI 0.01 to 0.05; P = 0.001; $I^2$ = 0%). Once again, the study by Vanaky et al[68] was excluded in this meta-analysis, due to its high risk of reporting bias. In a subsequent

| Study or subgroup | WBE Mean | SD | Total | LBE Mean | SD | Total | Weight | Mean difference IV, random, 95% CI | Mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **2.1.1 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | 0.034 | 0.084 | 8 | 0.091 | 0.084 | 9 | 4.3% | −0.06 [−0.14, 0.02 ] | |
| **Subtotal (95% CI)** | | | **8** | | | **9** | **4.3%** | **−0.06 [−0.14, 0.02]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 1.40$ ($P = 0.16$) | | | | | | | | | |
| | | | | | | | | | |
| **2.1.2 10 months** | | | | | | | | | |
| Murtezani et al[63] | 0.006 | 0.05 | 31 | 0.049 | 0.05 | 33 | 45.7% | −0.04 [−0.07, −0.02] | |
| **Subtotal (95% CI)** | | | **31** | | | **33** | **45.7%** | **−0.04 [−0.07, −0.02]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 3.44$ ($P = 0.0006$) | | | | | | | | | |
| | | | | | | | | | |
| **2.1.3 6 months** | | | | | | | | | |
| Kemper et al[61] | 0.005 | 0.084 | 15 | −0.0002 | 0.084 | 15 | 7.6% | 0.01 [−0.05, 0.07] | |
| Novaes et al[64] | 0.003 | 0.036 | 17 | 0.047 | 0.036 | 14 | 42.4% | −0.04 [−0.07, −0.02] | |
| **Subtotal (95% CI)** | | | **32** | | | **29** | **50.0%** | **−0.03 [−0.07, 0.02]** | |
| Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 2.18$, $df = 1$ ($P = 0.14$); $I^2 = 54\%$ | | | | | | | | | |
| Test for overall effect: $Z = 1.17$ ($P = 0.24$) | | | | | | | | | |
| | | | | | | | | | |
| **Total (95% CI)** | | | **71** | | | **71** | **100.0%** | **−0.04 [−0.06, 0.02]** | |
| Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 2.50$, $df = 3$ ($P = 0.48$); $I^2 = 0\%$ | | | | | | | | | |
| Test for overall effect: $Z = 4.77$ ($P < 0.00001$) | | | | | | | | | |
| Test for subgroup differences: $\chi^2 = 0.53$, $df = 2$ ($P = 0.77$); $I^2 = 0\%$ | | | | | | | | | |

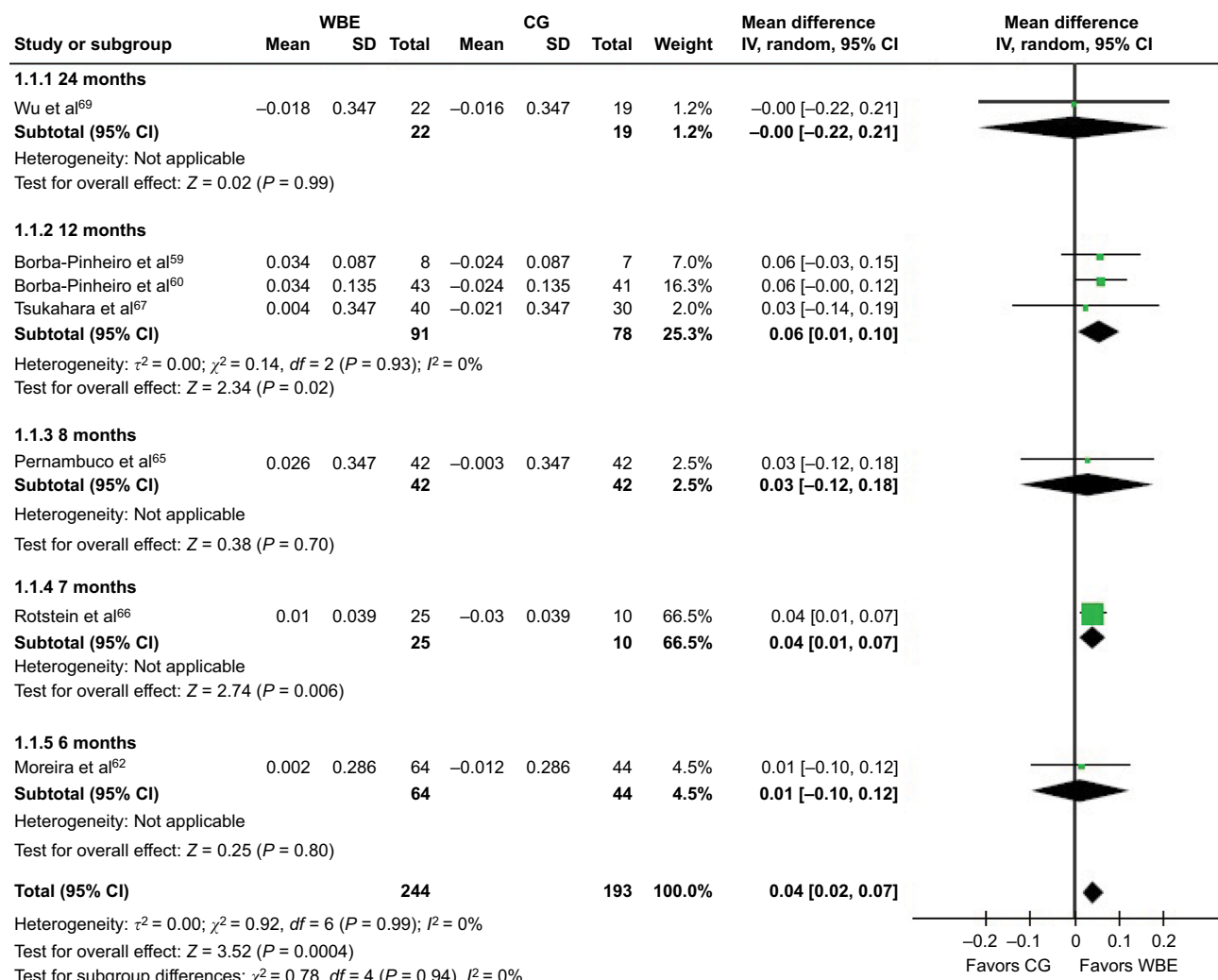−0.2 −0.1 0 0.1 0.2
Favors LBE    Favors WBE

**Figure 5** Forest plot of comparison between WBE and LBE for changes in LS BMD (mean difference in g/cm$^2$).
**Abbreviations:** WBE, water-based exercise; IV, inverse variance; LBE, land-based exercise; LS, lumbar spine; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.

| Study or subgroup | WBE Mean | SD | Total | CG Mean | SD | Total | Weight | Mean difference IV, random, 95% CI | Mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **1.2.1 24 months** | | | | | | | | | |
| Wu et al[69] | 0.029 | 0.049 | 22 | −0.002 | 0.049 | 19 | 37.1% | 0.03 [0.00, 0.06] | |
| **Subtotal (95% CI)** | | | **22** | | | **19** | **37.1%** | **0.03 [0.00, 0.06]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 2.02$ ($P = 0.04$) | | | | | | | | | |
| | | | | | | | | | |
| **1.2.2 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | −0.007 | 0.104 | 8 | −0.06 | 0.104 | 7 | 3.0% | 0.05 [−0.05, 0.16] | |
| Borba-Pinheiro et al[60] | −0.007 | 0.142 | 43 | −0.06 | 0.142 | 41 | 9.1% | 0.05 [−0.01, 0.11] | |
| **Subtotal (95% CI)** | | | **51** | | | **48** | **12.1%** | **0.05 [0.00, 0.11]** | |
| Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.00$, $df = 1$ ($P = 1.00$); $I^2 = 0\%$ | | | | | | | | | |
| Test for overall effect: $Z = 1.97$ ($P = 0.05$) | | | | | | | | | |
| | | | | | | | | | |
| **1.2.3 7 months** | | | | | | | | | |
| Rotstein et al[66] | 0.002 | 0.678 | 25 | −0.023 | 0.678 | 10 | 0.1% | 0.03 [−0.47, 0.52] | |
| **Subtotal (95% CI)** | | | **25** | | | **10** | **0.1%** | **0.03 [−0.47, 0.52]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 0.10$ ($P = 0.92$) | | | | | | | | | |
| | | | | | | | | | |
| **1.2.4 6 months** | | | | | | | | | |
| Moreira et al[62] | 0.01 | 0.067 | 64 | −0.014 | 0.067 | 44 | 50.7% | 0.02 [−0.00, 0.05] | |
| **Subtotal (95% CI)** | | | **64** | | | **44** | **50.7%** | **0.02 [−0.00, 0.05]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 1.83$ ($P = 0.07$) | | | | | | | | | |
| | | | | | | | | | |
| **Total (95% CI)** | | | **162** | | | **121** | **100.0%** | **0.03 [0.01, 0.05]** | |
| Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.95$, $df = 4$ ($P = 0.92$); $I^2 = 0\%$ | | | | | | | | | |
| Test for overall effect: $Z = 3.22$ ($P = 0.001$) | | | | | | | | | |
| Test for subgroup differences: $\chi^2 = 0.95$, $df = 3$ ($P = 0.81$); $I^2 = 0\%$ | | | | | | | | | |

−0.5 −0.25 0 0.25 0.5
Favors CG    Favors WBE

**Figure 6** Forest plot of comparison between WBE and CG for changes in FN BMD (mean difference in g/cm$^2$).
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; FN, femoral neck; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.

| Study or subgroup | WBE Mean | SD | Total | LBE Mean | SD | Total | Weight | Mean difference IV, random, 95% CI | Mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **2.2.1 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | −0.007 | 0.072 | 8 | 0.083 | 0.072 | 9 | 25.3% | −0.09 [−0.16, −0.02] | |
| **Subtotal (95% CI)** | | | **8** | | | **9** | **25.3%** | **−0.09 [−0.16, −0.02]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 2.57$ ($P = 0.01$) | | | | | | | | | |
| **2.2.2 6 months** | | | | | | | | | |
| Kemper et al[61] | 0.012 | 0.072 | 15 | −0.002 | 0.072 | 15 | 32.4% | 0.01 [−0.04, 0.07] | |
| Novaes et al[64] | 0 | 0.044 | 17 | −0.037 | 0.044 | 14 | 42.3% | −0.04 [−0.07, −0.01] | |
| **Subtotal (95% CI)** | | | **32** | | | **29** | **74.7%** | **−0.02 [−0.07, 0.03]** | |
| Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 2.76$, $df = 1$ ($P = 0.10$); $I^2 = 64\%$ | | | | | | | | | |
| Test for overall effect: $Z = 0.63$ ($P = 0.53$) | | | | | | | | | |
| **Total (95% CI)** | | | **40** | | | **38** | **100.0%** | **−0.03 [−0.08, 0.01]** | |
| Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 5.91$, $df = 2$ ($P = 0.05$); $I^2 = 66\%$ | | | | | | | | | |
| Test for overall effect: $Z = 1.37$ ($P = 0.17$) | | | | | | | | | |
| Test for subgroup differences: $\chi^2 = 2.97$, $df = 1$ ($P = 0.09$); $I^2 = 66.3\%$ | | | | | | | | | |

−0.2 −0.1 0 0.1 0.2
Favors LBE     Favors WBE

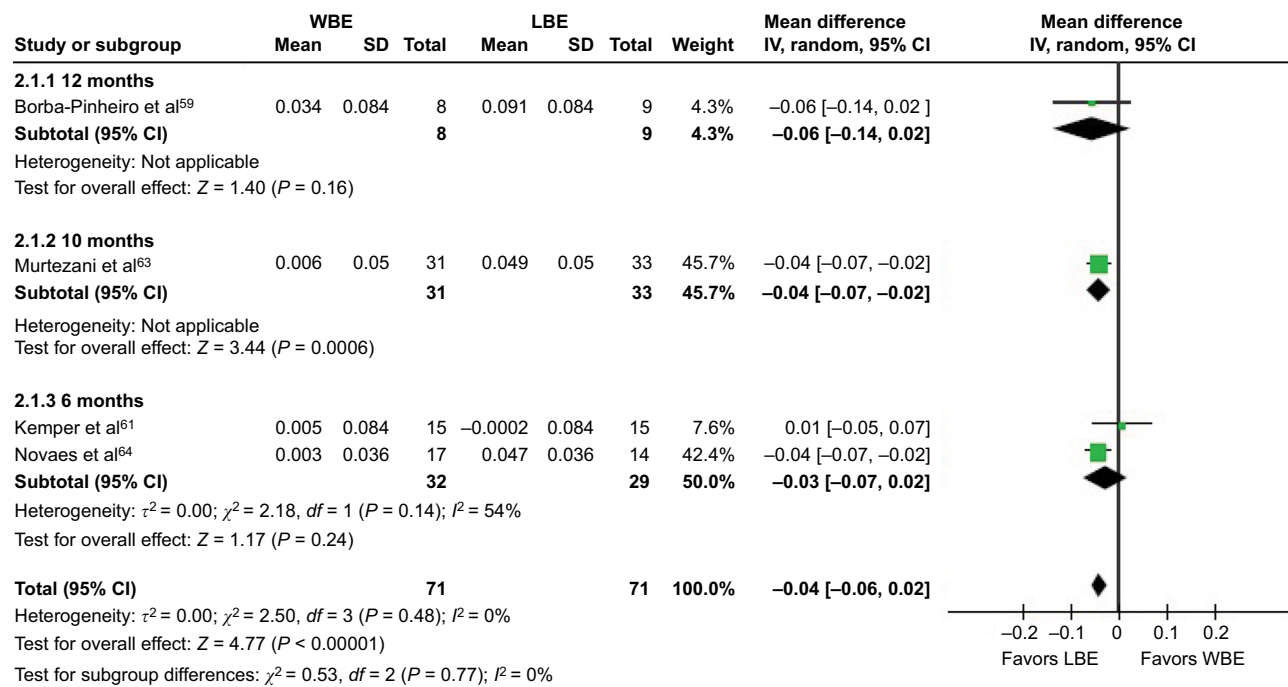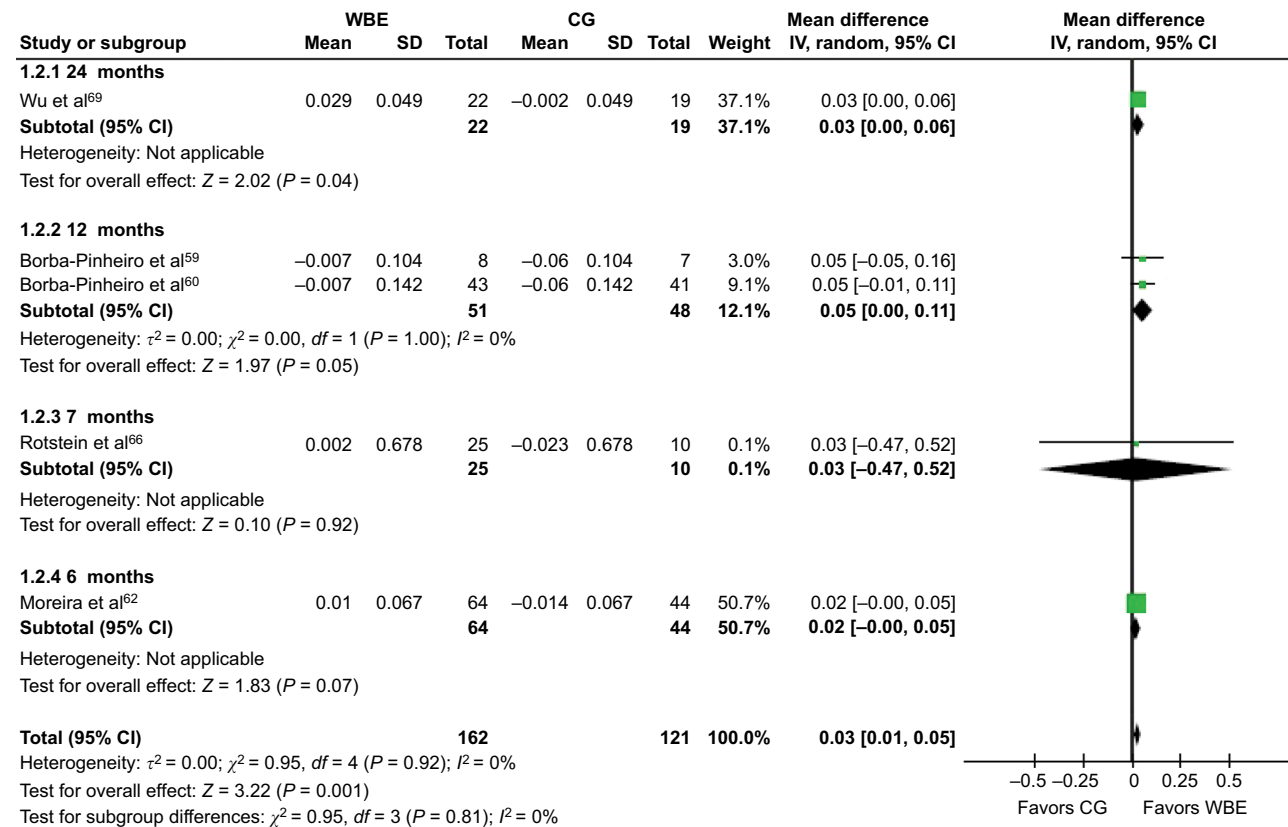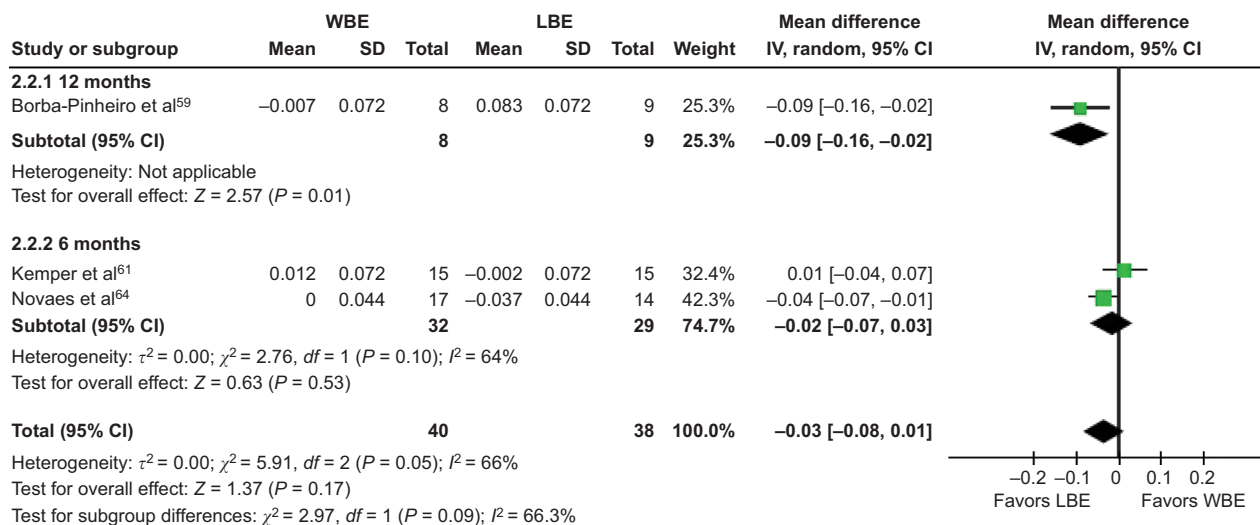**Figure 7** Forest plot of comparison between WBE and LBE for changes in FN BMD (mean difference in g/cm²).
**Abbreviations:** WBE, water-based exercise; IV, inverse variance; LBE, land-based exercise; FN, femoral neck; BMD, bone mineral density; CI, confidence interval.

sensitivity analysis, when this study was included, the results did not change, and heterogeneity was minimal ($I^2 = 14\%$, $P = 0.32$). In a further meta-analysis, there was no difference observed between WBE and LBE interventions in changes in FN BMD (MD −0.03 g/cm²; 95% CI −0.08 to 0.01; $P = 0.17$; $I^2 = 66\%$); however, heterogeneity was substantial ($P = 0.05$), as shown in Figure 7.

## GT BMD

Of the four studies that examined GT BMD, three reported an increase in GT BMD in the WBE group,[59,60,69] and the change was statistically significant in one study.[69] The fourth study did not report a change in the GT BMD value in the WBE group.[62] Four studies reported changes in GT BMD in the CG. Three reported a decrease,[59,60,62] which was statistically significant in one,[62] and one reported a nonsignificant increase.[69] Three studies reported a statistically significant difference between the WBE and CG groups in changes in GT BMD, in favor of WBE.[60,62,69] Only one study described a change in GT BMD in the LBE group, reporting an increase, but no reference was provided to the statistical significance of the result,[59] and when LBE was compared to WBE with regard to changes in GT BMD, no statistical difference was found between the two groups. Meta-analysis was conducted to compare the effects of WBE and CG on GT BMD. The results revealed a statistically significant difference in favor of WBE (MD 0.04 g/cm²; 95% CI 0.00 to 0.07; $P = 0.05$; $I^2 = 86\%$), as detailed in Figure 8. In order to address the considerable heterogeneity among studies in this particular meta-analysis, we conducted a sensitivity analysis to examine the impact of removing from the analysis the study conducted by Moreira et al,[62] and the results were still in favor of WBE,

with no heterogeneity then evident across the results (MD 0.05 g/cm²; 95% CI 0.03 to 0.07; $P < 0.00001$; $I^2 = 0\%$), as shown in Figure 9.

## WT BMD

Two studies assessed changes in WT BMD in the WBE group, and both reported a nonsignificant increase following WBE.[59,69] The same studies reported WT BMD results for a CG, and both described a nonsignificant decrease. A statistically significant difference between WBE and CG in their effects on WT BMD was observed, in favor of the WBE group, in the study conducted by Wu et al.[69] One of the studies also described a change in WT BMD for an LBE group, reporting a nonsignificant increase in that group, and no differences between WBE and LBE in their effects on WT BMD.[59] Meta-analysis revealed a significant difference between WBE and CG in their effects on WT BMD (MD 0.04 g/cm²; 95% CI 0.00 to 0.08; $P = 0.04$; $I^2 = 0\%$), as presented in Figure 10.

## TF BMD

Two studies described changes in TF BMD in a WBE group and a CG, reporting nonsignificant increases in TF BMD following WBE and nonsignificant decreases in TF BMD in the CG.[62,65] No significant differences were reported between these groups, and no significant differences were found in the results of a meta-analysis (MD 0.02 g/cm²; 95% CI −0.01 to 0.05; $P = 0.15$; $I^2 = 0\%$), as detailed in Figure 11.

## BMC

Only one study reported BMC as an outcome measure.[66] Change in BMC was described for both LS and FN, in both WBE and CG. The authors reported a nonstatistically

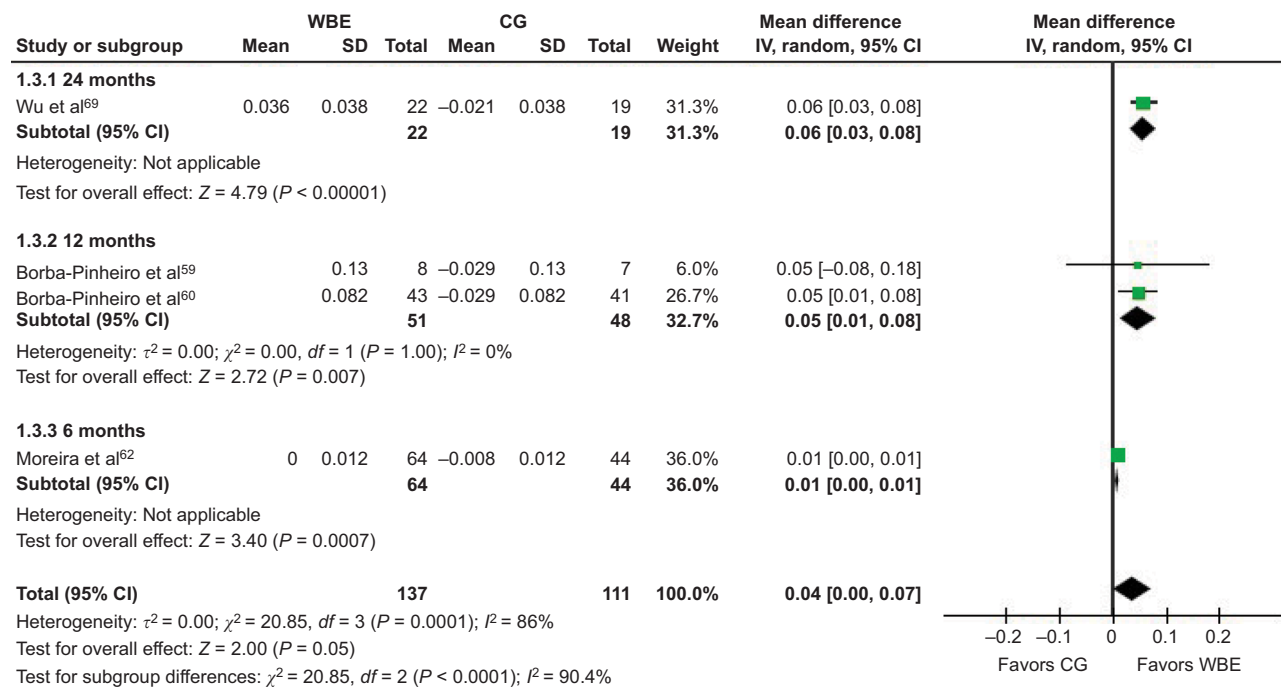| Study or subgroup | WBE Mean | SD | Total | CG Mean | SD | Total | Weight | Mean difference IV, random, 95% CI | Mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **1.3.1 24 months** | | | | | | | | | |
| Wu et al[69] | 0.036 | 0.038 | 22 | −0.021 | 0.038 | 19 | 31.3% | 0.06 [0.03, 0.08] | |
| **Subtotal (95% CI)** | | | **22** | | | **19** | **31.3%** | **0.06 [0.03, 0.08]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z$ = 4.79 ($P$ < 0.00001) | | | | | | | | | |
| **1.3.2 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | | 0.13 | 8 | −0.029 | 0.13 | 7 | 6.0% | 0.05 [−0.08, 0.18] | |
| Borba-Pinheiro et al[60] | | 0.082 | 43 | −0.029 | 0.082 | 41 | 26.7% | 0.05 [0.01, 0.08] | |
| **Subtotal (95% CI)** | | | **51** | | | **48** | **32.7%** | **0.05 [0.01, 0.08]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 0.00, $df$ = 1 ($P$ = 1.00); $I^2$ = 0% | | | | | | | | | |
| Test for overall effect: $Z$ = 2.72 ($P$ = 0.007) | | | | | | | | | |
| **1.3.3 6 months** | | | | | | | | | |
| Moreira et al[62] | 0 | 0.012 | 64 | −0.008 | 0.012 | 44 | 36.0% | 0.01 [0.00, 0.01] | |
| **Subtotal (95% CI)** | | | **64** | | | **44** | **36.0%** | **0.01 [0.00, 0.01]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z$ = 3.40 ($P$ = 0.0007) | | | | | | | | | |
| **Total (95% CI)** | | | **137** | | | **111** | **100.0%** | **0.04 [0.00, 0.07]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 20.85, $df$ = 3 ($P$ = 0.0001); $I^2$ = 86% | | | | | | | | | |
| Test for overall effect: $Z$ = 2.00 ($P$ = 0.05) | | | | | | | | | |
| Test for subgroup differences: $\chi^2$ = 20.85, $df$ = 2 ($P$ < 0.0001); $I^2$ = 90.4% | | | | | | | | | |

−0.2 −0.1 0 0.1 0.2
Favors CG    Favors WBE

**Figure 8** Forest plot of comparison between WBE and CG for changes in GT BMD (mean difference in g/cm$^2$).
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; GT, great trochanter; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.
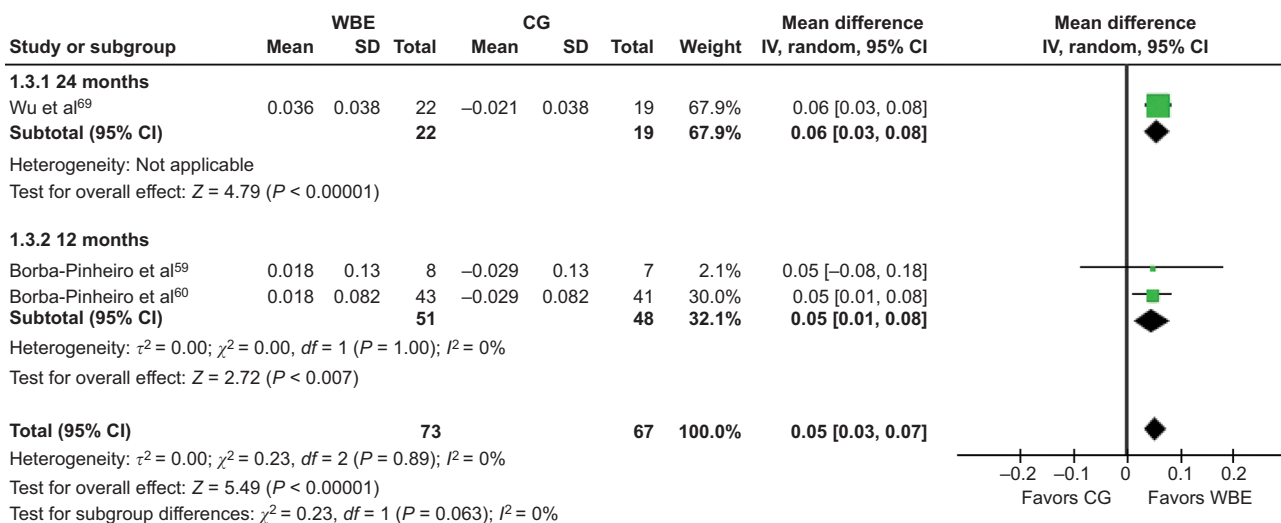
| Study or subgroup | WBE Mean | SD | Total | CG Mean | SD | Total | Weight | Mean difference IV, random, 95% CI | Mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **1.3.1 24 months** | | | | | | | | | |
| Wu et al[69] | 0.036 | 0.038 | 22 | −0.021 | 0.038 | 19 | 67.9% | 0.06 [0.03, 0.08] | |
| **Subtotal (95% CI)** | | | **22** | | | **19** | **67.9%** | **0.06 [0.03, 0.08]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z$ = 4.79 ($P$ < 0.00001) | | | | | | | | | |
| **1.3.2 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | 0.018 | 0.13 | 8 | −0.029 | 0.13 | 7 | 2.1% | 0.05 [−0.08, 0.18] | |
| Borba-Pinheiro et al[60] | 0.018 | 0.082 | 43 | −0.029 | 0.082 | 41 | 30.0% | 0.05 [0.01, 0.08] | |
| **Subtotal (95% CI)** | | | **51** | | | **48** | **32.1%** | **0.05 [0.01, 0.08]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 0.00, $df$ = 1 ($P$ = 1.00); $I^2$ = 0% | | | | | | | | | |
| Test for overall effect: $Z$ = 2.72 ($P$ < 0.007) | | | | | | | | | |
| **Total (95% CI)** | | | **73** | | | **67** | **100.0%** | **0.05 [0.03, 0.07]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 0.23, $df$ = 2 ($P$ = 0.89); $I^2$ = 0% | | | | | | | | | |
| Test for overall effect: $Z$ = 5.49 ($P$ < 0.00001) | | | | | | | | | |
| Test for subgroup differences: $\chi^2$ = 0.23, $df$ = 1 ($P$ = 0.063); $I^2$ = 0% | | | | | | | | | |

−0.2 −0.1 0 0.1 0.2
Favors CG    Favors WBE

**Figure 9** Forest plot of comparison between WBE and CG for changes in GT BMD (Moreira et al[62] excluded; mean difference in g/cm$^2$).
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; GT, great trochanter; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.

significant increase in BMC at both these sites in the WBE group and a nonstatistically significant decrease in BMC at both sites in the CG group. In the comparison between these groups, both LS and FN BMC increased significantly more in the WBE group than in the CG.

## Bone metabolism

Two studies included bone metabolism as an outcome measure,[62,65] and both compared the WBE results to results of a CG. Moreira et al[62] analyzed the biomarker of bone formation, procollagen type 1 amino-terminal propeptide (P1NP),

and the biomarker of bone resorption, carboxy-terminal cross-linking telopeptide of type 1 collagen (CTx), comparing the effects of WBE and CG on these biomarkers. The authors reported a mean increase in P1NP in both groups; however, the increase was statistically significant only in the WBE group. In the comparison between groups for P1NP, the effect on P1NP was significantly greater in the WBE group. The bone resorption biomarker CTx was observed to increase in both WBE and CG, but this increase reached statistical significance only in the CG, and no differences were found between these groups in their effects on CTx.
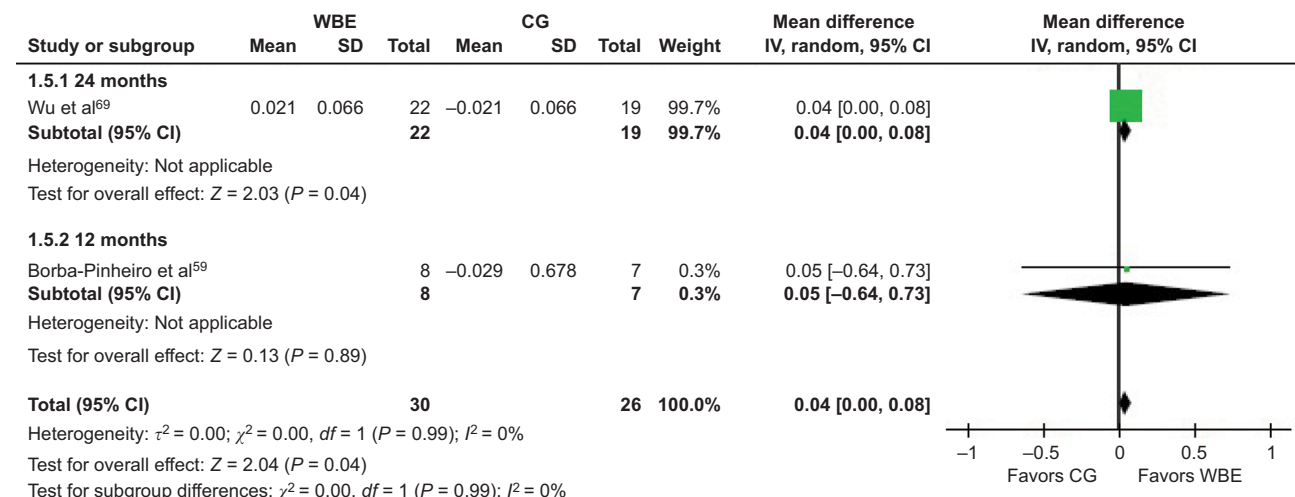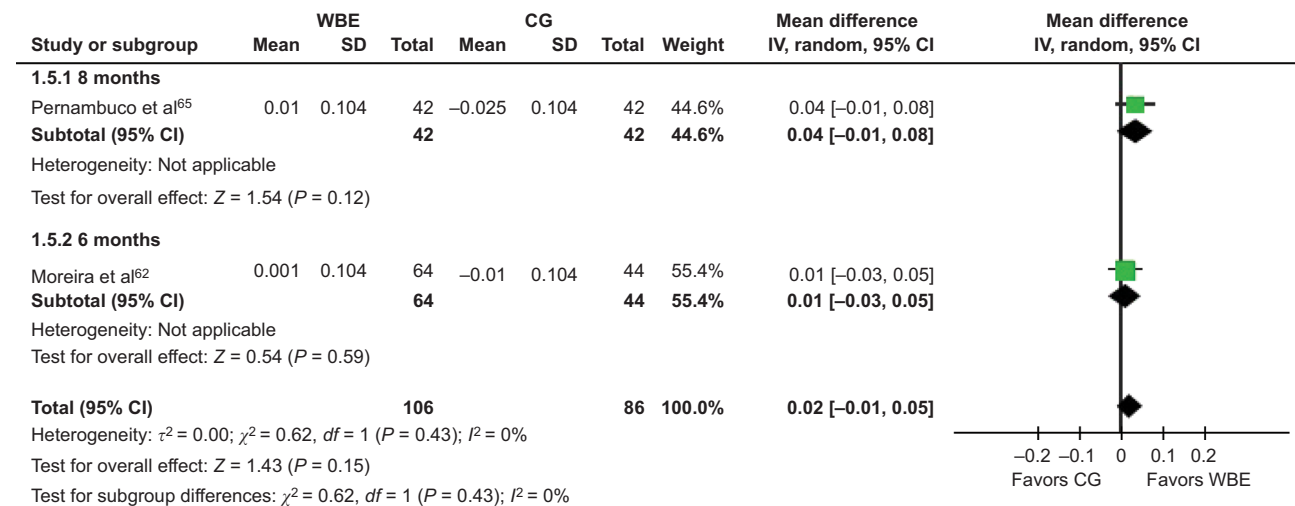
| | WBE | | | CG | | | | Mean difference | Mean difference |
|---|---|---|---|---|---|---|---|---|---|
| Study or subgroup | Mean | SD | Total | Mean | SD | Total | Weight | IV, random, 95% CI | IV, random, 95% CI |
| **1.5.1 24 months** | | | | | | | | | |
| Wu et al[69] | 0.021 | 0.066 | 22 | −0.021 | 0.066 | 19 | 99.7% | 0.04 [0.00, 0.08] | |
| **Subtotal (95% CI)** | | | **22** | | | **19** | **99.7%** | **0.04 [0.00, 0.08]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 2.03$ ($P = 0.04$) | | | | | | | | | |
| **1.5.2 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | | | 8 | −0.029 | 0.678 | 7 | 0.3% | 0.05 [−0.64, 0.73] | |
| **Subtotal (95% CI)** | | | **8** | | | **7** | **0.3%** | **0.05 [−0.64, 0.73]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 0.13$ ($P = 0.89$) | | | | | | | | | |
| **Total (95% CI)** | | | **30** | | | **26** | **100.0%** | **0.04 [0.00, 0.08]** | |

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.00$, $df = 1$ ($P = 0.99$); $I^2 = 0\%$
Test for overall effect: $Z = 2.04$ ($P = 0.04$)
Test for subgroup differences: $\chi^2 = 0.00$, $df = 1$ ($P = 0.99$); $I^2 = 0\%$



**Figure 10** Forest plot of comparison between WBE and CG for changes in WT BMD (mean difference in g/cm²).
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; WT, Ward's triangle; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.

| | WBE | | | CG | | | | Mean difference | Mean difference |
|---|---|---|---|---|---|---|---|---|---|
| Study or subgroup | Mean | SD | Total | Mean | SD | Total | Weight | IV, random, 95% CI | IV, random, 95% CI |
| **1.5.1 8 months** | | | | | | | | | |
| Pernambuco et al[65] | 0.01 | 0.104 | 42 | −0.025 | 0.104 | 42 | 44.6% | 0.04 [−0.01, 0.08] | |
| **Subtotal (95% CI)** | | | **42** | | | **42** | **44.6%** | **0.04 [−0.01, 0.08]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 1.54$ ($P = 0.12$) | | | | | | | | | |
| **1.5.2 6 months** | | | | | | | | | |
| Moreira et al[62] | 0.001 | 0.104 | 64 | −0.01 | 0.104 | 44 | 55.4% | 0.01 [−0.03, 0.05] | |
| **Subtotal (95% CI)** | | | **64** | | | **44** | **55.4%** | **0.01 [−0.03, 0.05]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: $Z = 0.54$ ($P = 0.59$) | | | | | | | | | |
| **Total (95% CI)** | | | **106** | | | **86** | **100.0%** | **0.02 [−0.01, 0.05]** | |

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.62$, $df = 1$ ($P = 0.43$); $I^2 = 0\%$
Test for overall effect: $Z = 1.43$ ($P = 0.15$)
Test for subgroup differences: $\chi^2 = 0.62$, $df = 1$ ($P = 0.43$); $I^2 = 0\%$



**Figure 11** Forest plot of comparison between WBE and CG for changes in TF BMD (mean difference in g/cm²).
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; TF, total femur; BMD, bone mineral density; CI, confidence interval; SD, standard deviation.

In the study conducted by Pernambuco et al,[65] the authors analyzed the biomarker of bone formation, osteocalcin. They reported a statistically significant increase in osteocalcin levels in the WBE group and a nonsignificant decrease in the CG. The mean increase in osteocalcin levels following WBE was significantly greater than that in the CG. Meta-analysis revealed significant differences between WBE and CG in favor of WBE for changes in the biomarkers of bone formation (SMD 0.49; 95% CI 0.20 to 0.78; $P = 0.0008$; $I^2 = 0\%$), as presented in Figure 12.

## Adverse events

Only three of the included studies reported information about adverse events. One of the studies reported that neither WBE group nor LBE group participants experienced fractures or serious orthopedic problems.[63] In that study,[63] one individual

allocated to the LBE group withdrew due to injury; however, it is not clear if the injury was associated with the exercise intervention. In another study, it was reported that no injuries were experienced by the participants in the WBE group.[62] In the study conducted by Kemper et al,[61] one individual was excluded due to chest pain during the WBE sessions. None of these three studies included fracture rate as an outcome, and no other study reported data regarding adverse events.

## Secondary outcomes
### Muscle strength

Two studies assessed muscle strength as an outcome.[63,69] Murtezani et al[63] assessed right-hand grip strength (GS), and right quadriceps strength (QS), and compared WBE to LBE. Both groups improved significantly in grip strength and QS; however, observed improvements following LBE were
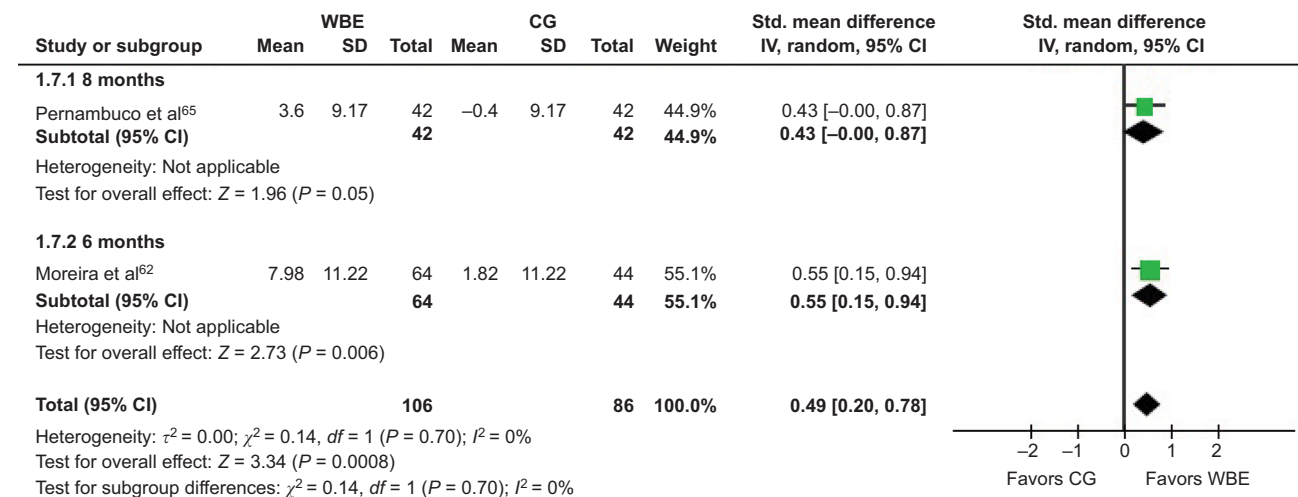
| Study or subgroup | WBE Mean | SD | Total | CG Mean | SD | Total | Weight | Std. mean difference IV, random, 95% CI | Std. mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **1.7.1 8 months** | | | | | | | | | |
| Pernambuco et al[65] | 3.6 | 9.17 | 42 | −0.4 | 9.17 | 42 | 44.9% | 0.43 [−0.00, 0.87] | |
| **Subtotal (95% CI)** | | | **42** | | | **42** | **44.9%** | **0.43 [−0.00, 0.87]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 1.96 (P = 0.05) | | | | | | | | | |
| | | | | | | | | | |
| **1.7.2 6 months** | | | | | | | | | |
| Moreira et al[62] | 7.98 | 11.22 | 64 | 1.82 | 11.22 | 44 | 55.1% | 0.55 [0.15, 0.94] | |
| **Subtotal (95% CI)** | | | **64** | | | **44** | **55.1%** | **0.55 [0.15, 0.94]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 2.73 (P = 0.006) | | | | | | | | | |
| | | | | | | | | | |
| **Total (95% CI)** | | | **106** | | | **86** | **100.0%** | **0.49 [0.20, 0.78]** | |
| Heterogeneity: $\tau^2$ = 0.00; $\chi^2$ = 0.14, df = 1 (P = 0.70); $I^2$ = 0% | | | | | | | | | |
| Test for overall effect: Z = 3.34 (P = 0.0008) | | | | | | | | | |
| Test for subgroup differences: $\chi^2$ = 0.14, df = 1 (P = 0.70); $I^2$ = 0% | | | | | | | | | |

−2 −1 0 1 2
Favors CG    Favors WBE

**Figure 12** Forest plot of comparison between WBE and CG for changes in biomarkers of bone formation.
**Abbreviations:** WBE, water-based exercise; CG, control group; IV, inverse variance; CI, confidence interval; SD, standard deviation.

significantly greater than those observed following WBE, for both outcomes. Wu et al[69] reported QS changes, comparing the results of WBE to a CG. The WBE was associated with a statistically significant increase in QS, whereas the CG was associated with a nonsignificant decrease, with no information provided about the level of statistical difference in this outcome between groups.

### Flexibility

No studies provided data on changes in participant flexibility associated with WBE. One study reported flexibility as an outcome,[63] using the "bend reach performance test". This study compared the WBE group to an LBE group, and the authors reported a statistically significant improvement in flexibility in the LBE group; however, no results were reported for the WBE group.

### Balance

Balance outcomes of participants were reported in two studies.[59,63] Both studies reported balance results for WBE and LBE, and one also provided results for a CG.[59] Borba-Pinheiro et al[59] assessed body balance using the Static Balance Test with Visual Control. Both WBE and LBE groups improved in their balance ability following the respective type of exercise, and the CG group decreased in balance ability; however, no information regarding statistical significance of these changes in balance within groups was reported. When the balance results of the WBE group were compared to those for LBE and CG, the differences in balance outcomes were not statistically significant. Murtezani et al[63] assessed balance using the Berg Balance Scale and reported positive changes in balance following WBE and LBE, which reached

statistical significance for the latter; however, no differences in balance outcomes were found between the WBE and LBE groups. Meta-analysis was conducted to compare effects of WBE and LBE on balance outcomes, and no statistically significant difference was found between the interventions (SMD −0.31; 95% CI −0.75 to 0.13; P = 0.17; $I^2$ = 0%), as detailed in Figure 13.

### Compliance

Only two studies reported levels of exercise compliance for both WBE and LBE groups.[61,64] Kemper et al[61] described an attendance rate of >75% of the sessions in both WBE and LBE groups, and Novaes et al[64] described an attendance rate of >85% of the sessions in both groups.

## Discussion

The main goal of the present systematic review was to determine the effects of WBE on bone health of middle-aged and older adults and to compare these WBE effects to those observed in a sedentary CG or LBE group. To the best of our knowledge, this is the first systematic review and meta-analysis addressing this topic. The main finding of the present systematic review supports the hypothesis that WBE may reduce age-related bone deterioration, as we identified statistically significant differences between WBE and CG in their effects on bone health, in favor of WBE. At the same time, the analyses also substantiate the belief that LBE is more effective than WBE in promoting positive changes in the bone tissue.

The importance of this review lies in the fact that medical and health/fitness professionals should be able to provide recommendations regarding effective alternatives among
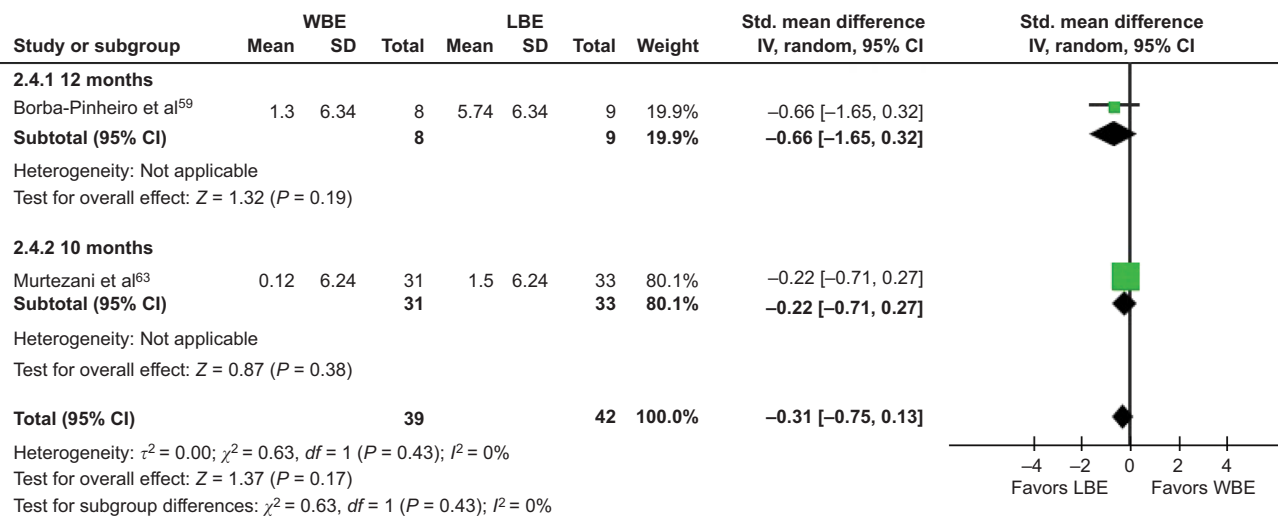
| Study or subgroup | WBE Mean | SD | Total | LBE Mean | SD | Total | Weight | Std. mean difference IV, random, 95% CI | Std. mean difference IV, random, 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| **2.4.1 12 months** | | | | | | | | | |
| Borba-Pinheiro et al[59] | 1.3 | 6.34 | 8 | 5.74 | 6.34 | 9 | 19.9% | −0.66 [−1.65, 0.32] | |
| **Subtotal (95% CI)** | | | 8 | | | 9 | 19.9% | **−0.66 [−1.65, 0.32]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 1.32 (P = 0.19) | | | | | | | | | |
| **2.4.2 10 months** | | | | | | | | | |
| Murtezani et al[63] | 0.12 | 6.24 | 31 | 1.5 | 6.24 | 33 | 80.1% | −0.22 [−0.71, 0.27] | |
| **Subtotal (95% CI)** | | | 31 | | | 33 | 80.1% | **−0.22 [−0.71, 0.27]** | |
| Heterogeneity: Not applicable | | | | | | | | | |
| Test for overall effect: Z = 0.87 (P = 0.38) | | | | | | | | | |
| **Total (95% CI)** | | | 39 | | | 42 | 100.0% | **−0.31 [−0.75, 0.13]** | |

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.63$, df = 1 (P = 0.43); $I^2 = 0\%$
Test for overall effect: Z = 1.37 (P = 0.17)
Test for subgroup differences: $\chi^2 = 0.63$, df = 1 (P = 0.43); $I^2 = 0\%$

−4 −2 0 2 4
Favors LBE    Favors WBE

**Figure 13** Forest plot of comparison between WBE and LBE for changes in body balance.
**Abbreviations:** WBE, water-based exercise; IV, inverse variance; LBE, land-based exercise; CI, confidence interval; SD, standard deviation.

exercise interventions, in order to keep the population physically active, preventing the bone loss associated with the aging process and subsequent increased risk of fracture. The findings of this review are consistent with findings of the systematic review conducted by Gomez-Bruton et al,[52] which investigated the effects of swimming in different age groups and revealed that WBE may have a positive impact on bone health in later adulthood. However, that review was limited to swimming, and the authors also concluded that the participants in the WBE had lower BMD than participants in land-based sports.

In the meta-analysis reported in this review comparing the effects of WBE to those of CG on LS BMD, the study conducted by Wu et al[69] was the only study that reported bone loss in both WBE and CG groups at this clinical site (Figure 4). However, the decrease in BMD described in that study did not reach statistical significance within either group, and no significant difference was evident between groups. As described in Table 1, the type of WBE used in this QE[69] was swimming, and the intervention was conducted 1.5 times per week for 24 months, with no information included pertaining to the duration and intensity of the sessions. It is, therefore, impossible to ascertain the actual dose of swimming received by participants, which may have been too low to have an effect on bone metabolism. As can be seen from Figure 4, the study having the greatest weight in this particular meta-analysis was the study conducted by Rotstein et al.[66] The WBE in that QE was hydrogymnastics, conducted three times per week for 7 months, in sessions of moderate to vigorous intensity, each lasting 60 minutes and involving participants who were postmenopausal women with normal BMD (Table 1). Thus, it could be argued that interventions

lasting >6 months, with sessions of similar intensity and duration to those described by Rotstein et al,[66] are likely to have positive effects on LS BMD. As shown in Figure 4, two RCTs were included in this particular meta-analysis focused on comparing the effects of WBE and CG on LS BMD,[62,65] and due to relative study weightings, these two RCTs contributed just 2.5% and 4.5% of the overall effect determined by the meta-analysis. The minor contribution of these RCTs is attributed to the relatively high SD associated with each. These values were obtained indirectly for both these RCTs, as we could not obtain SD values from the reported results, and contact with authors was not successful. As we chose a conservative approach to estimate the SD, the real value might be lower than the one used in our analysis, and this would influence the impact of each study on the outcome of the meta-analysis, but not the overall observed effect. The same interpretation applies to the small contribution of the study conducted by Wu et al[69] in this particular meta-analysis.

In the meta-analysis comparing the effects of WBE and LBE on LS BMD, it is worth noting that all four studies included in the meta-analysis reported a nonsignificant increase in LS BMD in the WBE group, and three reported a statistically significant increase in LS BMD in the LBE group (Figure 5). Surprisingly, the RCT conducted by Kemper et al[61] reported a nonsignificant decrease in LS BMD in the LBE group, which performed resistance training as the LBE intervention, while swimming was the WBE intervention (Table 1). The LBE and WBE sessions were conducted three times per week, in moderate to vigorous sessions of 60 minutes, for 6 months. The dose of the swimming intervention may explain the difference between the results reported by Kemper

et al[61] and by Wu et al[69] for swimming as a type of WBE – where Kemper et al[61] observed a nonsignificant increase in LS BMD following the swimming intervention, Wu et al,[69] who used a possibly much lower dose of swimming, observed a nonsignificant decrease. As depicted in Figure 5, two studies contributed with similar impact to this meta-analysis comparing effects of WBE and LBE on LS BMD, with respective weightings of 45.7% and 42.4% in the meta-analysis, attributed to their relatively small SD for this outcome.[63,64] The first was an RCT conducted by Murtezani et al[63] over 10 months, in which women with low BMD who were prescribed alendronate sodium and vitamin D were recruited (Table 1). The authors reported statistically significant differences between the groups in the observed changes in LS BMD, in favor of LBE; however, the differences reported might be explained by the fact that the exercise sessions were more frequent and lasted longer for individuals in the LBE group, with this LBE group, therefore, receiving a higher dose of exercise. The other study was conducted by Novaes et al[64] and was a QE conducted over 6 months, with exercise occurring three times per week, in sessions of 45 minutes of moderate to vigorous intensity (Table 1). These authors also reported statistically significant differences between the WBE and LBE groups in favor of LBE. Of note in the comparison between WBE and LBE with regard to their effects on LS BMD is that the study by Borba-Pinheiro et al[59] is the only study in which the WBE intervention was conducted for >6 months, and the WBE involved sessions of moderate to vigorous intensity lasting 60 minutes. In that study, there was no statistically significant difference observed between the groups in changes in LS BMD, and this finding might be explained by the small sample size, which also influenced the study's minor contribution to the overall effect observed in the meta-analysis (Figure 5).

In the comparison between WBE and CG with regard to their effects on FN BMD, the study conducted by Moreira et al[62] had a weighting of 50.7% in the meta-analysis, as a consequence of the small SD for this outcome measure (Figure 6). As detailed in Table 1, this RCT was conducted over 6 months, analyzing the effects of hydrogymnastics on bone health of previously sedentary women who were prescribed calcium and vitamin D. The WBE sessions were of moderate to vigorous intensity, conducted three times a week and lasted between 50 and 60 minutes. This study did not detect a statistically significant difference between WBE and CG in their effects on FN BMD, and a possible explanation for this finding might be the fact that the intervention was limited to 6 months. In the comparison of the effects of WBE and LBE

on FN BMD, the results are limited by substantial heterogeneity, as shown in Figure 7. The QE conducted by Novaes et al[64] contributed with a weighting of 42.3% to this meta-analysis, and the authors reported a statistically significant difference between the WBE and LBE groups, in favor of the LBE group. The exercise sessions of both groups lasted 45 minutes and involved moderate to vigorous intensity exercise, three times per week, and follow-up was limited to 6 months (Table 1). The RCT conducted by Kemper et al[61] contributed to increase the heterogeneity in the assessment of the overall effect of WBE when compared to LBE in this particular meta-analysis, as this study had contradictory results when compared to the other two studies included in the meta-analysis. The exercise sessions lasted 60 minutes and involved moderate to vigorous intensity exercise, conducted three times per week for 6 months, and the authors reported a nonsignificant increase in FN BMD in the WBE group and a nonsignificant decrease in the LBE group, with no differences found between the groups. Once again, these findings are consistent with the notion that WBE interventions conducted for a period of >6 months, in sessions of at least 60 minutes of moderate to high intensity and conducted three times per week, could possibly have a benefit to bone health. This hypothesis is also supported by the results reported for GT (Figures 8 and 9), WT (Figure 10), and TF (Figure 11); however, it was only possible to compare WBE to CG in the analysis of these three clinical sites.

Interestingly, Moreira et al[62] reported that both WBE and CG participants had a statistically significant increase in the biomarker of bone resorption CTx, although no differences were found between these groups. Those authors reported that levels of CTx typically increase in initial stages of the postmenopausal period, which was the case for the participants included in both groups. As shown in Figure 13, the RCT conducted by Murtezani et al[63] contributed with a weighting of 80% in the meta-analysis comparing the effects of WBE and LBE on balance ability, with this weighting being a consequence of the large sample size in that study. However, it is important to note that in that study the LBE group engaged in more frequent and longer exercise sessions than the WBE group. For measures of muscle strength and flexibility, no meta-analyses were conducted due to lack of studies reporting comparisons of these outcomes. The study of Murtezani et al[63] was the only one to report statistically significant differences between WBE and LBE for both these outcomes, each in favor of LBE, but once again it is important to highlight the differences between the LBE and WBE interventions used in this study, in terms of the frequency and duration of the exercise sessions discussed earlier. The findings of the

present review regarding effects of WBE on muscle strength and balance ability are in line with results of previous studies, which have demonstrated that individuals participating in WBE achieved a statistically significant improvement in both outcomes.[45,70–72] The studies conducted by Bergamin et al[70] and Oh et al[71] also reported a statistically significant improvement in flexibility for participants in WBE.

Only three studies included in the present review reported information regarding adverse events; however, due to lack of adequate reporting, no definitive conclusions can be drawn in this regard.

One of the strengths of this review is the comprehensive search of published studies, which was not limited by language of publication. This allowed us to include in the analyses two studies published in languages other than English, eliminating language bias in the review. For the meta-analyses, we used the random-effects model, as this enabled the researchers to estimate the mean effect across a range of studies in a manner that meant none of the individual studies could overly influence the overall estimate of effect. However, limitations should be highlighted. The generally low quality of available studies and the inclusion of QE in the meta-analyses mean that the results should be interpreted with caution. Another limitation is that none of the included studies reported the SD for the mean change in BMD. This value is necessary in order to conduct meta-analyses of the results, and so this value was estimated for each group. This estimate was derived for each study by either calculating the SD based on the reported *P*-value or by imputing the largest SD for that specific outcome that was reported in other studies. This approach was decided in order to achieve more conservative results but may have therefore also limited some of the effect sizes estimated in the meta-analyses. No study investigating a male population was found or included in this review, and so further research involving male participants is needed. It should also be noted that this review was purposely limited to investigating effects of WBE on bone health of middle-aged and older adults, and so the results should not be extrapolated to younger populations. While it is possible that publication bias may have affected the findings of this review, inspection of funnel plots developed to assess this likelihood suggested that it probably had little impact on the results.

## Conclusion

The results of this study corroborate the widely held belief that WBE is not as effective as LBE for enhancing bone health but they also indicate that, when the exercise dose is sufficient, WBE is better for bone health than a sedentary lifestyle in middle-aged and older adults. In order to increase exercise participation in middle-aged and older adults, it may be important to focus on alternative modes of exercise that are both suitable and feasible for this population, and which take into account possible clinical limitations of the individual and personal preferences. The results of the present meta-analyses indicate that an adequate dose of WBE may be a useful alternative to LBE, as it appears to decrease the rate of age-related bone loss in postmenopausal women. Moreover, it can increase BMD in this population, and it was demonstrated to have positive impacts on both bone metabolism and muscle strength.

There is currently not sufficient evidence to form a basis for recommending any specific WBE intervention when aiming to improve bone health; however, the results of this review suggest that WBE of higher intensity, frequency, and session duration, sustained over many months, is likely to be the most beneficial. Importantly, the findings of this review cannot be extrapolated to a male population since all participants in included studies were postmenopausal women, and they should not be extrapolated to younger populations, since the review was designed to focus only on middle-aged and older participants.

Further well-designed RCTs, including both males and females, should be undertaken to investigate the effects of WBE on bone health of middle-aged and older adults and to compare the effects of different types of WBE. Based on our findings, it appears that future interventions should be designed to last at least 12 months, and that the WBE sessions should be of moderate to vigorous intensity and at least 60 minutes in duration, occurring at least three times a week. With respect to BMD results, future research should adequately report SDs for the mean change within groups in this outcome measure, along with its *P*-value, in order to enable correct interpretation of the effect size of the results.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. NIH Consensus Development Panel on Osteoporosis Prevention D, and Therapy. Osteoporosis prevention, diagnosis, and therapy. *JAMA*. 2001;285(6):785–795.
2. Kanis JA, Melton LJ 3rd, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. *J Bone Miner Res*. 1994;9(8):1137–1141.
3. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet*. 2002;359(9319):1761–1767.
4. Papaioannou A, Morin S, Cheung AM, et al. 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary. *CMAJ*. 2010;182(17):1864–1873.

5. Henry MJ, Pasco JA, Nicholson GC, Kotowicz MA. Prevalence of osteoporosis in Australian men and women: Geelong Osteoporosis Study. *Med J Aust*. 2011;195(6):321–322.

6. Osteoporosis Australia Medical & Scientific Advisory Committee. What you need to know about osteoporosis: medical guide. *Osteoporosis Australia* [medical brochure]; 2014. Available from: http://www.osteoporosis.org.au/sites/default/files/files/oa_medical_brochure_2nd_ed.pdf. Accessed February 21, 2017.

7. Australian Institute of Health and Welfare (AIHW). Australia's health 2014. Australia's Health Series no. 14. Cat. no. AUS 178; 2014. Canberra: AIHW. Available from: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129548150. Accessed February 21, 2017.

8. Watts JJ, Abimanyi-Ochom J, Sanders KM. Osteoporosis Costing All Australians: a new burden of disease analysis – 2012 to 2022. *Osteoporosis Australia* [serial online]. Available from: http://osteoporosis.org.au/sites/default/files/files/Burden%20of%20Disease%20Analysis%202012-2022.pdf. Accessed February 21, 2017.

9. Bessette L, Ste-Marie LG, Jean S, et al. The care gap in diagnosis and treatment of women with a fragility fracture. *Osteoporos Int*. 2008;19(1):79–86.

10. Jones G, Nguyen T, Sambrook PN, Kelly PJ, Gilbert C, Eisman JA. Symptomatic fracture incidence in elderly men and women: the Dubbo Osteoporosis Epidemiology Study (DOES). *Osteoporos Int*. 1994; 4(5):277–282.

11. Bertram M, Norman R, Kemp L, Vos T. Review of the long-term disability associated with hip fractures. *Inj Prev*. 2011;17(6):365–370.

12. LeBlanc ES, Hillier TA, Pedula KL, et al. Hip fracture and increased short-term but not long-term mortality in healthy older women. *Arch Intern Med*. 2011;171(20):1831–1837.

13. Jiang HX, Majumdar SR, Dick DA, et al. Development and initial validation of a risk score for predicting in-hospital and 1-year mortality in patients with hip fractures. *J Bone Miner Res*. 2005;20(3):494–500.

14. Nazrun AS, Tzar MN, Mokhtar SA, Mohamed IN. A systematic review of the outcomes of osteoporotic fracture patients after hospital discharge: morbidity, subsequent fractures, and mortality. *Ther Clin Risk Manag*. 2014;10:937–948.

15. Laurent M, Gielen E, Claessens F, Boonen S, Vanderschueren D. Osteoporosis in older men: recent advances in pathophysiology and treatment. *Best Pract Res Clin Endocrinol Metab*. 2013;27(4): 527–539.

16. Bartl R. *Osteoporosis: Diagnosis, Prevention, Therapy*. 2nd rev. ed./in collaboration with Christoph Bartl ed. Berlin: Springer; 2009.

17. Haentjens P, Johnell O, Kanis JA, et al; Network on Male Osteoporosis in Europe (NEMO). Evidence from data searches and life-table analyses for gender-related differences in absolute risk of hip fracture after Colles' or spine fracture: Colles' fracture as an early and sensitive marker of skeletal fragility in white men. *J Bone Miner Res*. 2004;19(12):1933–1944.

18. Seeman E, Hopper JL, Young NR, Formica C, Goss P, Tsalamandris C. Do genetic factors explain associations between muscle strength, lean mass, and bone density? A twin study. *Am J Physiol*. 1996;270(2 pt 1): E320–E327.

19. Mikkola TM, Sipila S, Rantanen T, et al. Genetic and environmental influence on structural strength of weight-bearing and non-weight-bearing bone: a twin study. *J Bone Miner Res*. 2008;23(4):492–498.

20. Mikkola TM, Sipila S, Rantanen T, et al. Muscle cross-sectional area and structural bone strength share genetic and environmental effects in older women. *J Bone Miner Res*. 2009;24(2):338–345.

21. Rizzoli R, Bonjour JP, Ferrari SL. Osteoporosis, genetics and hormones. *J Mol Endocrinol*. 2001;26(2):79–94.

22. Kelley GA, Kelley KS, Kohrt WM. Exercise and bone mineral density in premenopausal women: a meta-analysis of randomized controlled trials. *Int J Endocrinol*. 2013;2013:741639.

23. Iolascon G, Resmini G, Tarantino U. Mechanobiology of bone. *Aging Clin Exp Res*. 2013;25(1):3–7.

24. Chen JH, Liu C, You L, Simmons CA. Boning up on Wolff's Law: mechanical regulation of the cells that make and maintain bone. *J Biomech*. 2010;43(1):108–118.

25. Prentice A. The relative contribution of diet and genotype to bone development. *Proc Nutr Soc*. 2001;60(1):45–52.

26. Zerwekh JE, Ruml LA, Gottschalk F, Pak CY. The effects of twelve weeks of bed rest on bone histology, biochemical markers of bone turnover, and calcium homeostasis in eleven normal subjects. *J Bone Miner Res*. 1998;13(10):1594–1601.

27. Khosla S, Atkinson EJ, Riggs BL, Melton LJ 3rd. Relationship between body composition and bone mass in women. *J Bone Miner Res*. 1996; 11(6):857–863.

28. Proctor DN, Melton LJ, Khosla S, Crowson CS, O'Connor MK, Riggs BL. Relative influence of physical activity, muscle mass and strength on bone density. *Osteoporos Int*. 2000;11(11):944–952.

29. Nilsson M, Ohlsson C, Mellstrom D, Lorentzon M. Sport-specific association between exercise loading and the density, geometry, and microstructure of weight-bearing bone in young adult men. *Osteoporos Int*. 2013;24(5):1613–1622.

30. Rubin CT, Lanyon LE. Regulation of bone formation by applied dynamic loads. *J Bone Joint Surg Am*. 1984;66(3):397–402.

31. Frost HM. The mechanostat: a proposed pathogenic mechanism of osteoporoses and the bone mass effects of mechanical and nonmechanical agents. *Bone Miner*. 1987;2(2):73–85.

32. Biewener AA. Safety factors in bone strength. *Calcif Tissue Int*. 1993;53(suppl 1):S68–S74.

33. Lanyon LE, Rubin CT. Static vs dynamic loads as an influence on bone remodelling. *J Biomech*. 1984;17(12):897–905.

34. Meunier PJ. Calcium, vitamin D and vitamin K in the prevention of fractures due to osteoporosis. *Osteoporos Int*. 1999;9(suppl 2):S48–S52.

35. World Health Organization. Prevention and management of osteoporosis. *World Health Organ Tech Rep Ser*. 2003;921:1–164,backcover.

36. Meléndez-Ortega A. Osteoporosis, falls and exercise. *Eur Rev Aging Phys Act*. 2007;4(2):61–70.

37. Gomez-Cabello A, Ara I, Gonzalez-Aguero A, Casajus JA, Vicente-Rodriguez G. Effects of training on bone mass in older adults: a systematic review. *Sports Med*. 2012;42(4):301–325.

38. Taaffe DR, Robinson TL, Snow CM, Marcus R. High-impact exercise promotes bone gain in well-trained female athletes. *J Bone Miner Res*. 1997;12(2):255–260.

39. Robling AG, Hinant FM, Burr DB, Turner CH. Shorter, more frequent mechanical loading sessions enhance bone mass. *Med Sci Sports Exerc*. 2002;34(2):196–202.

40. Yamazaki S, Ichimura S, Iwamoto J, Takeda T, Toyama Y. Effect of walking exercise on bone metabolism in postmenopausal women with osteopenia/osteoporosis. *J Bone Miner Metab*. 2004;22(5):500–508.

41. Chien MY, Wu YT, Hsu AT, Yang RS, Lai JS. Efficacy of a 24-week aerobic exercise program for osteopenic postmenopausal women. *Calcif Tissue Int*. 2000;67(6):443–448.

42. Bemben DA, Bemben MG. Dose-response effect of 40 weeks of resistance training on bone mineral density in older adults. *Osteoporos Int*. 2011;22(1):179–186.

43. Vainionpaa A, Korpelainen R, Leppaluoto J, Jamsa T. Effects of high-impact exercise on bone mineral density: a randomized controlled trial in premenopausal women. *Osteoporos Int*. 2005;16(2):191–197.

44. Hinman RS, Heywood SE, Day AR. Aquatic physical therapy for hip and knee osteoarthritis: results of a single-blind randomized controlled trial. *Phys Ther*. 2007;87(1):32–43.

45. Tsourlou T, Benik A, Dipla K, Zafeiridis A, Kellis S. The effects of a twenty-four-week aquatic training program on muscular strength performance in healthy elderly women. *J Strength Cond Res*. 2006;20(4):811–818.

46. Meredith-Jones K, Waters D, Legge M, Jones L. Upright water-based exercise to improve cardiovascular and metabolic health: a qualitative review. *Complement Ther Med*. 2011;19(2):93–103.

47. Chevutschi A, Lensel G, Vaast D, Thevenon A. An electromyographic study of human gait both in water and on dry ground. *J Physiol Anthropol*. 2007;26(4):467–473.

48. Bellew JW, Gehrig L. A comparison of bone mineral density in adolescent female swimmers, soccer players, and weight lifters. *Pediatr Phys Ther*. 2006;18(1):19–22.

49. Taaffe DR, Snow-Harter C, Connolly DA, Robinson TL, Brown MD, Marcus R. Differential effects of swimming versus weight-bearing activity on bone mineral status of eumenorrheic athletes. *J Bone Miner Res*. 1995;10(4):586–593.

50. Velez NF, Zhang A, Stone B, Perera S, Miller M, Greenspan SL. The effect of moderate impact exercise on skeletal integrity in master athletes. *Osteoporos Int*. 2008;19(10):1457–1464.

51. Balsamo S, Mota LM, Santana FS, et al. Resistance training versus weight-bearing aquatic exercise: a cross-sectional analysis of bone mineral density in postmenopausal women. *Rev Bras Reumatol*. 2013;53(2):193–198.

52. Gomez-Bruton A, Gonzalez-Aguero A, Gomez-Cabello A, Casajus JA, Vicente-Rodriguez G. Is bone tissue really affected by swimming? A systematic review. *PLoS One*. 2013;8(8):e70119.

53. Higgins JPT, Green S, editors [homepage on the Internet]. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org/. Accessed February 21, 2017.

54. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–269,W264.

55. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev*. 2012;1(1):1–9.

56. Thomson Reuters. *EndNote* [computer program]. Version X7.2.1 (Bld 8311). New York, NY: Thomson Reuters; 2016.

57. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011; 343:d5928.

58. The Nordic Cochrane Centre, The Cochrane Collaboration. *Review Manager (RevMan)* [computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2014.

59. Borba-Pinheiro CJ, de Alencar Carvalho MC, da Silva NS, Drigo AJ, Bezerra JC, Dantas EH. Bone density, balance and quality of life of postmenopausal women taking alendronate participating in different physical activity programs. *Ther Adv Musculoskelet Dis*. 2010;2(4):175–185.

60. Borba-Pinheiro C, Gurgel de Alencar Carvalho M, Drigo A, Lima da Silva N, Pernambuco C, Dantas E. Bone variables after a water aerobics program in postmenopausal women undergoing pharmacological treatment. *Sport Sci Health*. 2012;7(1):13–18.

61. Kemper C, de Oliveira RJ, Bottaro M, et al. Effects of swimming and resistance training on bone mineral density of older women. *Revista Brasileira de Medicina do Esporte*. 2009;15(1):10–13.

62. Moreira LD, Fronza FC, Dos Santos RN, et al. The benefits of a high-intensity aquatic exercise program (HydrOS) for bone metabolism and bone mass of postmenopausal women. *J Bone Miner Metab*. 2014; 32(4):411–419.

63. Murtezani A, Nevzati A, Ibraimi Z, Sllamniku S, Meka VS, Abazi N. The effect of land versus aquatic exercise program on bone mineral density and physical function in postmenopausal women with osteoporosis: a randomized controlled trial. *Ortop Traumatol Rehabil*. 2014;16(3):319–325.

64. Novaes GS, Novaes JS, Vilaca JA, et al. Effects of 24 weeks of strength training or hydrogymnastics on bone mineral density in postmenopausal women. *J Sports Med Phys Fitness*. 2013;53(suppl 1 to No. 3):51–56.

65. Pernambuco CS, Borba-Pinheiro CJ, Vale RG, Di Masi F, Monteiro PK, Dantas EH. Functional autonomy, bone mineral density (BMD) and serum osteocalcin levels in older female participants of an aquatic exercise program (AAG). *Arch Gerontol Geriatr*. 2013;56(3): 466–471.

66. Rotstein A, Harush M, Vaisman N. The effect of a water exercise program on bone density of postmenopausal women. *J Sports Med Phys Fitness*. 2008;48(3):352–359.

67. Tsukahara N, Toda A, Goto J, Ezawa I. Cross-sectional and longitudinal studies on the effect of water exercise in controlling bone loss in Japanese postmenopausal women. *J Nutr Sci Vitaminol (Tokyo)*. 1994; 40(1):37–47.

68. Vanaky B, Sadeghi H, Ramezani N. The effect of 12 weeks weight bearing water training on the bone density of middle age sedentary women. *Biosci Biotechnol Res Asia*. 2014;11(2):931–936.

69. Wu J, Yamakawa J, Tabata I, Yoshitake Y, Higuchi M. The effects of swimming exercise on bone mineral density in postmenopausal women. *Jpn J Phys Fit Sport*. 2000;49(5):543–548.

70. Bergamin M, Ermolao A, Tolomio S, Berton L, Sergi G, Zaccaria M. Water- versus land-based exercise in elderly subjects: effects on physical performance and body composition. *Clin Interv Aging*. 2013;8: 1109–1117.

71. Oh S, Lim JM, Kim Y, Kim M, Song W, Yoon B. Comparison of the effects of water- and land-based exercises on the physical function and quality of life in community-dwelling elderly people with history of falling: a single-blind, randomized controlled trial. *Arch Gerontol Geriatr*. 2015;60(2):288–293.

72. Bento PC, Pereira G, Ugrinowitsch C, Rodacki AL. The effects of a water-based exercise program on strength and functionality of older adults. *J Aging Phys Act*. 2012;20(4):469–483.