



Effects of dynamic resistance exercise on bone mineral density in postmenopausal women: a systematic review and meta-analysis with special emphasis on exercise parameters

M. Shojaa¹ · S. von Stengel¹ · M. Kohl² · D. Schoene¹ · W. Kemmler¹

Received: 27 January 2020 / Accepted: 28 April 2020 / Published online: 12 May 2020
© The Author(s) 2020

Abstract

This systematic review and meta-analysis set out to determine the effect of dynamic resistance exercise (DRT) on areal bone mineral density (aBMD) in postmenopausal women and derive evidence-based recommendations for optimized training protocols. A systematic review of the literature according to the PRISMA statement included (a) controlled trials, (b) of isolated DRT with at least one exercise and one control group, (c) with intervention durations ≥ 6 months, (d) aBMD assessments at lumbar spine or proximal femur, (e) in cohorts of postmenopausal women. We searched eight electronic databases up to March 2019 without language restrictions. The meta-analysis was performed using a random-effects model. Standardized mean differences (SMD) for BMD changes at lumbar spine (LS), femoral neck (FN), and total hip (TH) were defined as outcome measures. Moderators of the exercise effects, i.e., “intervention length,” “type of DRT,” “training frequency,” “exercise intensity,” and “exercise volume,” were addressed by sub-group analyses. The study was registered in the international prospective register of systematic reviews (PROSPERO) under ID: CRD42018095097. Seventeen articles with 20 exercise and 18 control groups were eligible. SMD average is 0.54 (95% CI 0.22–0.87) for LS-BMD, 0.22 (0.07–0.38) for FN-BMD, and 0.48 (0.22–0.75) for TH-BMD changes (all $p \leq 0.015$). While sub-group analysis for FN-BMD revealed no differences within categories of moderators, lower training frequency (< 2 sessions/week) resulted in significantly higher BMD changes at LS and TH compared to higher training frequency (≥ 2 sessions/week). Additionally, free weight training was significantly superior to DRT devices for improving TH-BMD. This work provided further evidence for significant, albeit only low–moderate, effects of DRT on LS-, FN-, and TH-BMD. Unfortunately, sub-analysis results did not allow meaningful exercise recommendations to be derived. This systematic review and meta-analysis observed a significant low–moderate effect of dynamic resistance exercise on bone mineral density changes in postmenopausal women. However, sub-group analyses focusing on exercise characteristics found no results that enable the derivation of meaningful exercise recommendations in the area of exercise and osteoporosis prevention or therapy.

Keywords Bone mineral density · Dynamic resistance exercise · Exercise characteristics · Postmenopausal women

Introduction

Physical exercise is the most powerful non-pharmaceutical fracture prevention strategy in postmenopausal women [1, 2]. Dynamic resistance training (DRT) as defined as any kind of resistance exercise that involves joint movement and

focuses on the development of musculoskeletal strength is considered an important component of osteoporosis prevention and therapy [3–5]. Nevertheless, with respect to areal bone mineral density (aBMD), recent meta-analysis on DRT reported on average low exercise-induced effect sizes at best (review in [6–9]). Reviewing the underlying studies, however, there is some evidence that not only isolated DRT protocols were included in the meta-analysis. Further, large variations between the individual study findings can be observed. The main reason for this outcome can be attributed to the complexity of exercise interventions with respect to exercise variables (e.g., exercise intensity, duration, frequency), training principles (e.g., progression, periodization), and training conditions (e.g., supervision, devices) [10]. Previous meta-analysis on

✉ W. Kemmler
wolfgang.kemmler@imp.uni-erlangen.de

¹ Institute of Medical Physics, Friedrich-Alexander University Erlangen-Nürnberg, Henkestrasse 91, 91052 Erlangen, Germany

² Department of Medical and Life Sciences, University of Furtwangen, Schwenningen, Germany

DRT did not account for training parameters. At best, there was separate analysis for “exercise intensity” [6]. Thus, apart from meta-analysis that properly focuses on isolated DRT protocols, there is a need for research aimed at identifying effective training parameters to generate recommendations for exercise protocols in the area of osteoporosis.

Therefore, in the present systematic review and meta-analysis, our primary study aim was to determine the effect (size) of isolated DRT on BMD at lumbar spine and proximal femur regions of interest (ROI) in postmenopausal women in comparison with control groups and further (secondary study aim) to identify relevant exercise characteristics by sub-analysis to derive recommendations for optimized exercise protocols in clinical practice.

Material and methods

Data sources and search strategy

The present study on DRT was based on a comprehensive systematic review of the effect of exercise on (areal) BMD in postmenopausal women. This systematic review and meta-analysis followed strictly the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [11]. The study was registered in the international prospective register of systematic reviews (PROSPERO) (ID: CRD42018095097). Briefly, eight electronic databases (PubMed, Scopus, Web of Science, Cochrane, Science Direct, Eric, ProQuest, and Primo) were searched for relevant articles published up to March 1, 2019 without language restrictions. The search strategy related to the population, intervention, and outcomes and was constructed around search terms for “Bone Mineral Density”, “Exercise,” and “Postmenopausal.” Key words and their synonyms were used to sensitize the search by applying the following query: (“Bone” or “Bone mass” or “Bone status” or “Bone structure” or “Bone turnover” or “Bone metabolism” or “Bone mineral content” or “Skeleton” or “Bone Mineral Density” or “BMD” or “Bone Density” or “Osteoporosis” or “Osteoporosis” or “Osteopenia”) AND (“Postmenopause” or “Post-Menopause” or “Postmenopausal”) AND (“Exercise” or “Training” or “Athletic” or “Sport” or “physical activity”) AND (“Clinical trial” or “Randomized clinical trial”). Unpublished reports or congress abstracts were not considered. One reviewer (MS) conducted the search and, following the omission of duplicate publication, screened studies by title and abstracts according to the eligibility criteria. In the secondary search, reference lists of articles included were reviewed to further identify relevant and eligible studies.

In summary, 42 authors were contacted by e-mail due to the following issues: (a) participants/group allocation, (b) missing drop out data, (c) separation of the pre- or postmenopausal

group, (d) result given in a graph only, (e) no mean change of BMD given, (f) missing standard deviation (SD).

Inclusion and exclusion criteria

We included studies (or study arms) if they met the following criteria: (a) randomized and non-randomized controlled trials with at least one exercise group as an intervention versus one control group with sedentary/habitual active lifestyle without exercise or with sham exercise; (b) women with postmenopausal status at study onset; (c) intervention of at least 6 months; (d) areal BMD of the LS or/and the proximal femur regions “TH” and/or “FN” were listed as outcome measures at baseline and follow-up assessment; (e) BMD determined by dual-energy X-ray absorptiometry (DXA) or dual-photon absorptiometry (DPA); (f) $\leq 10\%$ of participants on hormone (replacement) therapy (HT or HRT), osteoanabolic/antiresorptive (e.g., bisphosphonate, denosumab, strontium ranelate), or osteocatabolic (glucocorticoids) and pharmaceutical agents, albeit only when the number of users was comparable between exercise and control. For the present analysis on isolated DRT effects, we further included only studies (f) that applied isolated DRT without any adjuvant exercise component and without bone specific warm-ups with validated effect on bone (e.g., running, hopping, aerobic dance).

We excluded studies that included (a) mixed gender or mixed pre- and postmenopausal cohorts without separate BMD analysis for postmenopausal women¹, (b) women undergoing chemo- and/or radiotherapy, and (c) women with diseases that relevantly affect bone metabolism. Further, (d) double/multiple publications from one study and preliminary data from subsequently published trials and (e) review articles, case reports, editorials, conference abstracts, and letters were not considered. For the present analysis on isolated DRT effects, we likewise excluded studies (f) that reported a pre-study history of RT type exercise ≥ 60 min/week 1 year prior to the study intervention.

Data extraction

Two reviewers (SvS and MS) independently evaluated full-text articles and extracted data from the included studies. If they could not reach a consensus, a third reviewer was consulted (KW). We designed a pre-piloted extraction form to extract relevant data. This covered the publication characteristics (e.g., (first) author’s name, title, country, and publication year), methodology (i.e., design, objectives, sample size for each group), participant characteristics (i.e., age, weight, BMI, years since menopause), exercise characteristics (e.g.,

¹ We do not exclude studies with participants that were not community dwelling

intervention duration, training frequency, exercise intensity, movement/velocity, progression), compliance/withdrawals, risk assessment, BMD, and outcome characteristics.

Outcome measures

The primary outcomes in the present study was the change in (areal) BMD at lumbar spine, total hip (TH), and femoral neck (FN) regions of interest (ROI) as assessed by DXA or DPA between baseline and follow-up.

Quality assessment

All the articles that satisfied the predefined inclusion criteria were independently assessed for risk of bias by two independent raters (WK and MV) using the PEDro (Physiotherapy Evidence Database) scale [12, 13]. Disagreements were solved by discussion including a third assessor (SvS) until a consensus was reached. We classified the methodological quality of the included studies as follows: ≥ 7 = high, 5–6 = moderate, and < 5 = low [14].

Data synthesis

According to the “Cochrane Handbook for Systematic Reviews of Interventions” [15], standard deviation (SD) can be obtained from the standard error (SE) or confidence interval (CI) by using the following formulas²:

$$SD = SE \times \sqrt{N}$$

$$SD = \sqrt{N} \times (\text{upper limit} - \text{lower limit}) / 3.92$$

Further, authors ($n = 11$) were contacted to provide missing SDs. In cases of no reply or unavailable data ($n = 11$), the exact p value of the absolute change in BMD was obtained to compute the SD of the mean change. In cases of unreported p values ($n = 11$), we calculated the SDs using pre- and post-SDs and correlation coefficients according to the Cochrane Handbook of Systemic Reviews [15]. Lastly, when the absolute mean difference was not available, it was imputed by calculation of the difference between post- and pre-intervention ($n = 7$). In cases of multiple BMD assessments, we considered only changes between the baseline and final BMD assessments.

To identify potential predictors of successful DRT protocols, we applied several sub-group analyses for the following: (a) intervention period (≤ 8 months vs. 9–18 months vs. > 18 months); (b) type of RT (machines vs. free weight vs. both types); (c) net training frequency³ (< 2 vs. ≥ 2 sessions/week); (d) exercise intensity (low ($< 65\%$ 1RM) vs. moderate ($65\text{--}80\%$ 1RM) vs. high ($\geq 80\%$ 1RM)); and (e) exercise volume

per session (exercises \times sets \times repetitions) as structured in low (< 160 reps/session), moderate (160 to < 300 reps session), and high (≥ 300 reps/session) volume.

Statistical analysis

The statistical analysis was performed using the statistical software R (R Development Core Team) [16]. Effect size (ES) value was considered as the standardized mean differences (SMDs) combined with the 95% confidence interval (CI).

Random-effects meta-analysis was performed by applying the metafor package [17]. Heterogeneity for between-study variability was determined using the Cochran Q test; comparable to other statistical analysis, a p value < 0.05 was considered significant. The level of heterogeneity was analyzed with the I^2 statistic. An I^2 of between 0 and 40% is considered as low, 30 to 60% as moderate, and 50 to 90% as substantial heterogeneity, respectively [15]. For those studies with two different intervention groups, the control group was proportionally split into two groups for comparison against each intervention group [15]. Sensitivity analysis was conducted to check whether the overall result of the analysis is robust regarding the use of the imputed correlation coefficient. Funnel plots with regression test and the rank correlation between effect estimates and their standard errors (SEs), using the t test and Kendall's τ statistic respectively, were applied to explore potential publication bias. To adjust the results for possible publication bias, we also conducted a trim and fill analysis using the L0 estimator proposed by Duval et al. [18]. A p value of < 0.05 was considered as significant for all tests.

In order to identify potential moderators of exercise, sub-group analyses were performed with the exercise parameters and their corresponding categories as listed above.

Results

Study characteristics and quality assessments

In total, our search identified 17 eligible studies ([19–35]; Fig. 1), with 20 exercise and 18 control groups. Table 1 displays participant baseline characteristics of the included studies.

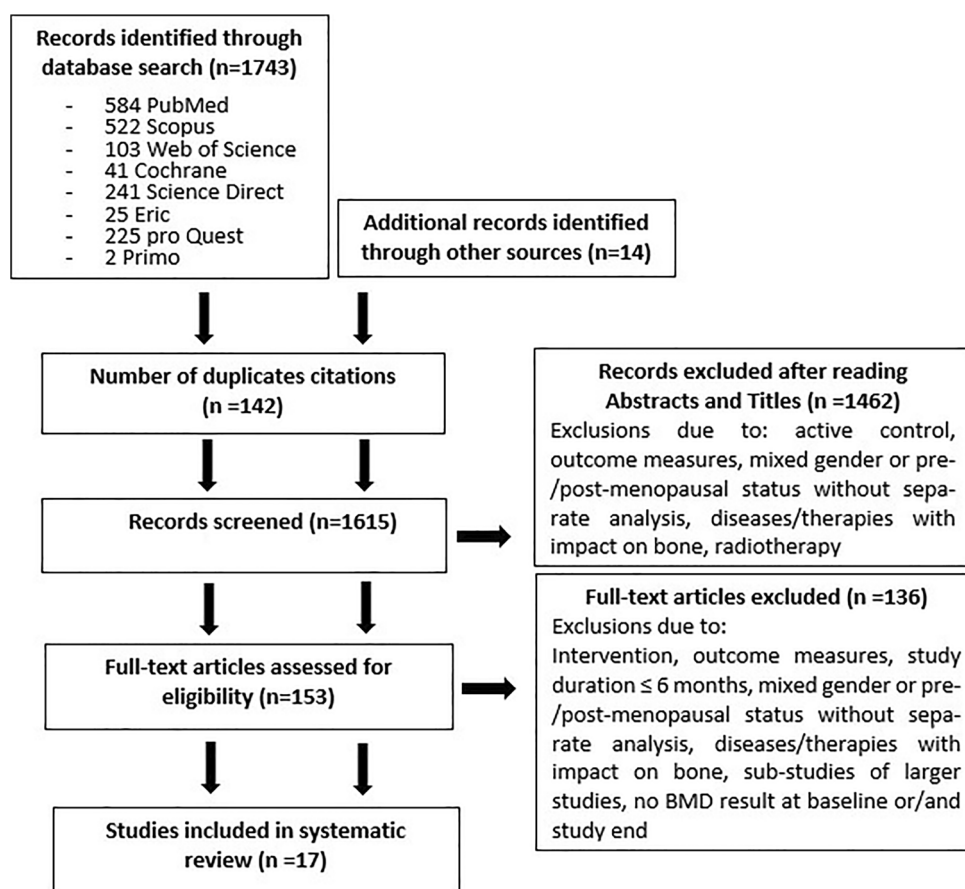
The pooled number of participants in the intervention and control groups was 423 and 373 women, respectively. Sample sizes in the exercise arms ranged from 10 [31] to 35 participants [27] per group. Only one study [24] included women with osteopenia/osteoporosis⁴; none of the other studies considered bone status (Table 2). Age of the postmenopausal women in the studies ranged between 41 and 60 years [19] and 65–82 years [32]. Menopausal age varied from 0.5 [34] to

² One study reported CI and 8 studies utilized SE.

³ ...considering participants attendance rate.

⁴ $> 30\%$ of bone tissue loss as determined by X-ray assessment at LS or hip.

Fig. 1 Flow diagram of search process according to PRISMA [11]



> 21 years post [26]. Three studies [19, 27, 31] focused particularly on cohorts of “early postmenopausal women” (1–≤ 7 years post). Average body mass index among the groups averaged from 23.1 [28] to 28.7 kg/m² [19]. Six studies included participants with sedentary/habitually active lifestyles or at least no prior RT exposure [20, 22, 26, 28, 30, 33, 35], 8 trials involved participants with exercise activities presumably with minor effects on bone [19, 21, 23–25, 29, 31, 32], and two studies did not provide corresponding information [27, 34]. The studies were conducted in Australia [29], Brazil [22, 30], Canada [21, 23, 36], Germany [24], Hong Kong [35], and the USA [19, 20, 25–28, 31, 32, 34].

Intervention characteristics

Vitamin-D and calcium supplementation

Six studies provided Vit-D and/or calcium supplementation [19, 21, 23, 26, 32] in their exercise and control groups.

Exercise

Table 2 specifies the exercise protocols of included studies. Apart from three studies [19, 25, 32], all trials compared a single exercise group (EG) with a single non-exercise control

group (CG). From the criteria, all the studies applied DRT; the majority of studies used conventional RT machines or free weights, one study [22] applied “Pilates exercises” with specific devices (e.g., reformer, cadillac), and another study [35] used an resistance band of low–moderate strength.

Length of the intervention (or study) ranged from six [19, 22, 24, 29] to 24 months [34]; no study reported a delay between the end of the intervention and the control assessment. Most studies focused on all or most main muscle groups.⁵ Kerr et al. [25] applied a unilateral DRT that stresses the “ipsilateral forearm and hip region,” Sinaki et al. [34] focused on back strengthening in a prone position, Woo et al. [35] predominately conducted hip and lower limb exercises, and Maddalozzo et al. [27] specified back squats and deadlifts. Apart from two studies that did not provide sufficient information for the LS-site [25, 35], all the studies addressed their specified BMD ROI (i.e., LS and hip) with specific exercises.

Most studies prescribed a training frequency of three sessions per week (Table 2); however, when considering participant attendance, the net training frequency of five studies [24, 26–29] was on average below two sessions per week. Length

⁵ In addition to the DRT, Kohrt et al. [26] applied 2–3 × 10 min bouts of intense rowing exercise.

Table 1 Baseline characteristics of participants in included studies ($n = 17$)

First author, year	Initial sample size (n)	Age (years)	Menopausal age (years)	Body mass (kg)	Height (cm)	BMI (kg/m ²)
Bemben ^a , 2000	HI: 13	HI: 50 ± 2	HI: 4 ± 1	HI: 74.7 ± 5.6	HI: 162 ± 2	HI: 28.7 ± 2.4
	HR: 11	HR: 52 ± 2	HR: 2 ± 1	HR: 62.7 ± 3.4	HR: 165 ± 2	HR: 23.2 ± 1.2
	C: 11	C: 52 ± 1	C: 3 ± 1	C: 66.5 ± 4.2	C: 166 ± 2	C: 24.2 ± 1.7
Bemben 2010	E: 22	E: 64 ± 1	> 5	E: 76.6 ± 3.2	E: 161 ± 2	E: 30 ± 1
	C: 12	C: 63 ± 1		C: 77.9 ± 4.5	C: 163 ± 1	C: 29 ± 1
Chilibeck, 2002 [*]	E: 14	E: 57 ± 2	E: 9 ± 2	E: 72 ± 4.3	E: 164 ± 2	E: 27.0 ± 1.7
	C: 14	C: 59 ± 2	C: 8 ± 2	C: 73.2 ± 4.8	C: 165 ± 1	C: 26.6 ± 1.2
de Oliveira, 2018	E: 17	E: 56 ± 7	E: 8 ± 7	E: 67.4 ± 8.6	E: 157 ± 6	E: 27.2 ± 2.7
	C: 17	C: 54 ± 5	C: 9 ± 7	C: 64.6 ± 6.6	C: 154 ± 4	C: 27.3 ± 2.5
Duff, 2016	E: 22	E: 65 ± 5	n.g.	n.g.	E: 162 ± 6	n.g.
	C: 22	C: 65 ± 5	n.g.	n.g.	C: 160 ± 7	n.g.
Hartard 1996	E: 18	E: 64 ± 6	E > 2	E: 67 ± 7.7	E: 162 ± 7	n.g.
	C: 16	C: 67 ± 10	C > 2	C: 63.8 ± 11.2	C: 158 ± 6	n.g.
Kerr ^b , 1996	HI: 28	HI: 58 ± 4	HI: 8 ± 3	HI: 69.4 ± 11.4	HI: 165 ± 7	n.g.
	HR: 28	HR: 56 ± 5	HR: 6 ± 4	HR: 70.8 ± 10	HR: 165 ± 6	n.g.
Kohrt, 1997 [*]	E: 15	E: 65 ± 1	n.g.	E: 72.6 ± 2.3	E: 164 ± 2	n.g.
	C: 15	C: 68 ± 1	n.g.	C: 71.6 ± 1.8	C: 163 ± 2	n.g.
Madda-lozzo, 2007	E: 35	E: 52 ± 3	E: 2 ± 1	E: 70 ± 8.7	n.g.	n.g.
	C: 34	C: 52 ± 3	C: 2 ± 1	C: 67.1 ± 12.6	n.g.	n.g.
Nelson, 1994	E: 21	E: 61 ± 4	E: 12 ± 5	E: 64.7 ± 7.7	E: 163 ± 6	E: 24.4 ± 2.5
	C: 19	C: 57 ± 6	C: 10 ± 5	C: 62.2 ± 8.9	C: 164 ± 8	C: 23.1 ± 2.2
Nicholson, 2015	E: 28	E: 66 ± 4	E: > 5	E: 70.6 ± 9.1	E: 164 ± 4	E: 26 ± 3.2
	C: 29	C: 66 ± 5	C: > 5	C: 66.8 ± 10.7	C: 163 ± 5	C: 24.5 ± 2.9
Orsatti, 2013	E + Pl: 20	E + Pl: 56 ± 9	E + Pl: 9 ± 6	n.g.	n.g.	E + Pl: 26 ± 3.0
	Pl: 20	Pl: 55 ± 8	Pl: 8 ± 6	n.g.	n.g.	Pl: 30.4 ± 5.3
Pruitt, 1992 [*]	E: 17	E: 54 ± 1	E: 3 ± 1	E: 64.2 ± 1.9	E: 162 ± 1	n.g.
	C: 10	C: 56 ± 1	C: 4 ± 1	C: 65.5 ± 2.9	C: 163 ± 2	n.g.
Pruitt, 1995	HI: 15	HI: 67 ± 1	n.g.	HI: 64.5 ± 9.2	HI: 162 ± 7	HI: 24.5 ± 3.4
	HR: 13	HR: 68 ± 1	n.g.	HR: 61.5 ± 4.6	HR: 160 ± 5	HR: 23.9 ± 1.6
	C: 12	C: 70 ± 4	n.g.	C: 63.8 ± 9.1	C: 160 ± 9	C: 25.1 ± 3.1
Rhodes, 2000	E: 22	E: 69 ± 3	n.g.	E: 68.4 ± 12	E: 161 ± 5	n.g.
	C: 22	C: 68 ± 3	n.g.	C: 61.7 ± 12.9	C: 159 ± 4	n.g.
Sinaki, 1989	E: 34	E: 56 ± 4	n.g.	E: 66.2 ± 9.3	E: 163 ± 6	n.g.
	C: 34	C: 56 ± 4	n.g.	C: 66.1 ± 10.6	C: 161 ± 5	n.g.
Woo, 2007	E: 30	E: 70 ± 3	n.g.	n.g.	n.g.	E: 24.6 ± 4.0
	C: 30	C: 69 ± 3	n.g.	n.g.	n.g.	C: 24.9 ± 3.0

All values are presented as mean ± SD, otherwise stated

HI high intensity, HR high repetition, C control (group), E exercise (group), Pl placebo, n.g. not given

*Values are presented as mean ± SE

^aBaseline data of the study completers ($n = 25$)

^bUnilateral loading (hip, forearm) with the contralateral side as control

of the exercise sessions varied from about 1–2 min (i.e., 10 × back extension [34]) to about 120 min (i.e., 36 sets × 20 reps, 2–3 min of rest between the sets). Most studies applied a multiple set approach ([19–21, 23–28, 30, 32, 33]. The protocol of Nicholson et al. [29] scheduled 10 × 4–6 min blocks of one (e.g., squats) or several exercises for the same muscle groups (e.g., chest, back, triceps). As a result, repetitions per set for a single exercise were up to 108 reps (132 reps/block) [29] for the latter study, but most studies applied sets with 7–12 repetitions [19–28,

30–34].⁶ Three studies [19, 25, 32], comparing high versus low intensity RT protocols, further scheduled sets of 14–20 reps in their low intensity study arms. Correspondingly, relative exercise intensity ranged between 80% 1RM [19, 20, 25, 26, 28, 30, 32] and ≤ 30% 1RM [29, 34]. Absolute exercise intensity (i.e., “effort”)

⁶ This relates to the high intensity exercise groups of three studies [19, 25, 32]. Further, one study [30] prescribed sets of 20–30 reps for two exercises.

Table 2 Exercise characteristics of included studies ($n = 16$)

Author, year	Health and exercise status	Length (months)	PR-INT	Type of exercise, amount of exercises, methods	Site specificity	Volume in min/w.; settings; (attendance rate)	Exercise/strain composition
Bemben, 2000 Low intensity	Healthy, no RT	6	Yes	DRT (most main muscle groups) on machines	LS: Yes TH: Yes	3 × 60; S-JE; (93%)	8 exercises, 3 sets, 16 reps, 40% 1RM
Bemben, 2000 High intensity	Healthy, no RT	6	Yes	DRT (most main muscle groups) on machines	LS: Yes TH: Yes	3 × 60; S-JE; (87%)	8 exercises, 3 sets, 8 reps, 80% 1RM
Bemben, 2010	Healthy, no RT	8	No	DRT (most main muscle groups) on machines	LS: Yes TH: Yes	3 × 60; S-JE; (92%)	8 exercises, 3 sets, 10 reps, 80% 1RM
Chilibeck, 2002	Healthy, no BSE	12	Yes	DRT (all main muscle groups) on machines	LS: Yes TH: Yes	3 × RT; S-JE; (78%)	12 exercises, 2 sets, 8–10 reps, ≈ 70–80% 1RM (“a weight that could be lifted comfortably for 10 reps”)
Campos de Oliveira, 2018	Healthy, Sed/HA	6	Yes	Pilates (strengthening and flexibility; all mains muscle groups) on Pilates “devices.”	LS: Yes TH: Yes	3 × 60; S-JE; (93%)	21 exercises (strengthening and flexibility), 1 set, 10 reps, RPE 5–6 at Borg CR10
Duff, 2016	Healthy, no RT	9	Yes	DRT (all main muscle groups) on machines and with free weights	LS: Yes TH: Yes	3 × RT; S-JE; (n.g.)	12 exercises, 2 sets of 8–12 reps to muscular fatigue
Hartard, 1996	Osteopenia, < 1 h week	6	Yes	DRT (all main muscle groups) on machines, stretching	LS: Yes TH: Yes	2 × ≈ 60; S-JE; n.g.	14 exercises, 2 sets, 8–12 reps to “marked fatigue” (≈ 70% 1RM), 3–4 s per reps
Kerr, 1996 Low intensity	Healthy, no RT	12	Yes	Unilateral DRT (randomized allocation of the upper or hip/lower limb to exercise or control) on machines or free weights.	LS:? TH: yes	3 × ≈ 120; S-JE; (82%)	12 exercises; 3 sets, 20 RM (≈ 60–65% 1RM), 2–3 min rest between sets
Kerr, 1996 High intensity	Healthy, no RT	12	Yes	Unilateral DRT (see above) on machines or free weights.	LS:? TH: yes	3 × ≈ 100 S-JE; (82%)	12 exercises; 3 sets, 8 RM (≈ 75–80% 1RM), 2–3 min rest between sets
Kohrt, 1997	Healthy, Sed/HA	11	Yes	DRT (most main muscle groups) on machines and with free weights, rowing	LS: Yes TH: Yes	5 × 45–50; n.g. (presumably S-JE); (≈ 70%)	DRT: 2 × week, 8 exercises, 2–3 sets, 8–12 reps “to fatigue” (≈ 70–80% 1RM) and rowing: 3 × week, 2–3 sets × 10 min at 80–85% HRmax
Maddalozzo, 2007	Healthy, n.g.	12	Yes	DRT (back squat, deadlifts) with free weights; subordinate: exercises that focus on alignment, flexibility, posture, abdominal strength	LS: Yes TH: Yes	2 × 50; S-JE; (85%)	2 exercises, 2 warm-ups sets, 10–12 reps, 50% 1RM; 3 sets, 8–12 reps, 60–75% 1RM; TUT: 1–2 s concentric – 0 s iso-metric – 2–3 s eccentric
Nelson, 1994	Healthy, Sed/HA	12	Yes	DRT (“most” main muscle groups) on machines.	LS: Yes TH: Yes	2 × 45; S-JE; (88%)	5 exercises, 3 sets, 8 reps, 80% 1RM; TUT-6-9 s/rep; 3 s rest between reps, 90–120 s rest between sets
Nicholson, 2015	Healthy, no RT	6	Yes	DRT (all main muscle groups): “Body Pump Release 83” (i.e., barbell exercises with (very) low intensity)	LS: Yes TH: Yes	2 × 50, S-JE; (89%)	10 × ≈ 4–6 min blocks of exercises for all main muscle groups (21 exercises in total); up to 108 reps (squats), ≤ 30% 1RM
Orsatti, 2013	Healthy, Sed/HA	9	Yes	DRT (all main muscle groups) on machines and with free weights	LS: Yes TH: Yes	3 × 50–60, S-JE; (n.g.)	8 exercises 3 sets at 8–12 RM; 3 sets, 20–30 reps for trunk flexion and calf raises; 1–2 min rest

Table 2 (continued)

Author, year	Health and exercise status	Length (months)	PR-INT	Type of exercise, amount of exercises, methods	Site specificity	Volume in min/w.; setting; (attendance rate)	Exercise/strain composition
Pruitt, 1992	Healthy, no BSE	9	Yes	DRT (all main muscle groups) on machines and with free weights	LS: Yes TH: Yes	3 × 60; S-JE; (83%)	11 exercises, 1 set at 10 RM (no more details given)
Pruitt, 1995 High intensity	Healthy no RT	12	Yes	DRT (all main muscle groups) on machines	LS: Yes TH: Yes	3 × 60; S-JE; (79%)	10 exercises, 1 warm up set, 14 reps, 40% and 2 sets, 7 reps, 80% IRM
Pruitt, 1995 Low intensity	Healthy No RT	12	Yes	DRT (all main muscle groups) on machines	LS: Yes TH: Yes	3 × 60; S-JE; (79%)	10 exercises, 3 sets, 14 reps, 40% IRM
Rhodes, 2000	Healthy, “no organized sports”	12	Yes	DRT (all main muscle groups) on machines	LS: Yes TH: Yes	3 × 60; PS-JE; (86%)	≥ 6 exercises (n.g. in detail), 3 sets, 8 reps, ≈ 75% IRM, TUT: 2–3 s concentric – 3–4 s eccentric movement/rep applied in a circuit mode.
Sinaki, 1989	Healthy, n.g.	24	No	DRT (back strengthening exercise in a prone position using a back pack; ≈ hyperextensions) with a free weight	LS: Yes -----	5 × ≈ 1–2 min; HE; n.g.	One back strengthening exercise, 1 set, 10 reps, with a weight „equivalent to 30% of the maximum isometric back muscle strength in pounds (maximum 23 kg)
Woo, 2007	Healthy, no RT	12	No	DRT (arm-lifting, hip-abduction, heel raise, hip-flexion,-extension, squat) using resistance bands	LS: ? TH: Yes	3 × ≈ 15 min; (n.g.); (76%)	6 exercises, 1 set, 30 reps with a resistance band “of medium strength” (no more information given).

Health status: We focused on reported osteoporosis/osteopenia and fractures only. Otherwise, women were listed as “healthy.” Exercise status: we mainly used the characterization of the authors. In some cases, we summarize the information given as “no bone-specific exercise.” Progression: We only considered the progression of exercise intensity during the intervention. Site specificity: Estimated site specificity of the exercise to address LS- or hip-BMD. Exercise volume/week, setting, attendance: Number of sessions per week × minutes per session (e.g., 3 × 60). Setting of the exercise session, i.e., either supervised group (S-JE), partially supervised individual (PS-JE) or home exercise/exercise individually performed without supervision (HE). Attendance defined as rate of sessions performed in relation to total exercise sessions (%). Composition of stram/exercise variables per session: number of exercises, number of sets, number of repetitions; exercise intensity, set endpoint (e.g., ... “to failure”), time under tension per rep/movement velocity, rest pauses between sets

BSE bone specific exercise, Sed/HA sedentary/habitually active, PrInt progression of exercise intensity, BSE bone-specific exercise, DRT dynamic resistance exercise, S supervised, PS partially supervised, JE joint exercise program, HE home exercise program, RPE rate of perceived exertion, TUT time under tension

was rarely prescribed, however. Five studies prescribed either work to repetition maximum [30, 31, 37] or work to muscular fatigue [23, 26] as a set endpoint [38]; another study [22] referred to 5–6 (i.e., strong–strong+) on the Borg CR10 scale. Reviewing the repetition number and relative exercise intensity (% 1RM), some studies [34] or study arms [19, 32] clearly exercised with low or very low effort.⁷ Time under load or velocity during the different sections of the movement (concentric–isometric–eccentric) [39] was also rarely mentioned [24, 27, 33]. Time under load varied from 3 to 4 s [40] to 5–6 s [36] per repetition. None of the studies reported an explosive movement in the concentric or eccentric phase. Apart from the 24 month study by Sinaki et al. [34], progression or at least regular adjustments of exercise intensity was realized by all the studies. The application of periodization models [41] was not reported by any of the studies.

Most studies focused on a supervised group exercise protocol ([19–25, 27–32, 35], while two other studies relied on partially supervised individual gym-based RT [33] or non-supervised exercise training at home [34].

The Pedro Score of the reviewed studies ranged from 4 to 8 of 10 total score points (Table 3). Level of agreement between the raters for methodological quality of the studies was 100%. The methodological quality of five studies can be considered as high [35],⁸ 10 were considered to be of moderate, and 2 of low quality. According to the mixed-effects analysis, no significant differences between the Pedro Score categories (high vs. moderate vs. low) were detected at LS- ($p = 0.639$), FN- ($p = 0.968$), or TH-BMD ($p = 0.416$).

Results for primary outcomes

Apart from two studies [31, 34] that applied DPA, all the others used DXA. Furthermore, all the other studies except two ([25]: hip only; [34]: LS only) determined both BMD at LS and proximal femur regions of interest.

Effect of exercise on LS-BMD

Sixteen studies with 18 exercise groups evaluated the effect of exercise on LS-BMD. In summary, the exercise interventions resulted in significant positive effects ($p = 0.001$). The pooled estimate of random effect analysis was SMD 0.54, 95% CI 0.22–0.87, however with a substantial level of heterogeneity between trials ($I^2 = 74.8\%$, $Q = 70.1$) (Fig. 2). A sensitivity analysis imputing minimum SD (best case; SMD 0.70, 95%

CI 0.27–1.13) or maximum SD (worst cases: SMD 0.42, 95% CI 0.17–0.67) resulted in significant result in all cases. We expect that the mean value imputation comes closest to the true effect (Fig. 2).

The funnel plot suggested positive evidence of publication bias (Fig. 3). The regression ($p = 0.920$) and rank ($p = 0.881$) correlation test for funnel plot asymmetry did not indicate any significant asymmetry. Adjusting for possible publication bias using a trim and fill analysis [18] did not result in varying results.

Sub-group analyses for LS-BMD

Intervention duration

Of 18 groups, 6 training groups conducted short-term interventions (≤ 8 months), 11 groups applied a moderate duration (9–18 months) intervention, and one training group scheduled a 24-month intervention (Table 2). According to a mixed-effects analysis, no significant difference was observed between the sub-groups ($p = 0.421$).

Type of exercise

Of 18 training groups, 10 groups worked with resistance training devices, four with free weights or resistance bands, and four conducted a mix of both types (Table 2). According to a mixed-effects analysis, no significant difference was observed between the sub-groups ($p = 0.700$).

Training frequency

Of 18 training groups, 5 groups exercised fewer than 2 sessions per week and 13 groups exercised ≥ 2 sessions per week (Table 2). According to a mixed-effects analysis, a significant difference was observed between the two groups ($p = 0.002$). The sub-group analysis demonstrated the highest effects sizes for the lower training frequency (SMD 1.26, 95% CI 0.88–1.64) compared to an SMD of 0.24 (95% CI -0.05 –0.54) in the sub-group that exercised ≥ 2 sessions per week.

Exercise intensity

Of 18 training groups, 5 groups exercised with low ($< 65\%$ 1RM), 7 with moderate (65 – $< 80\%$ 1RM), and 6 with high relative intensity ($\geq 80\%$ 1RM) (Table 2). According to a mixed-effects analysis, no significant difference was observed between the sub-groups ($p = 0.404$).

Exercise volume/session (exercises \times sets \times reps)

Of 18 training groups, 7 groups applied a low (< 160 reps/session), 8 a moderate (160 to < 300 reps/session), and

⁷ E.g. 10 reps at 30% maximum isometric back muscle strength [34] or 14–16 reps at 40% 1RM [19, 32]. The same might be true for some exercises of the protocol of Woo et al. [35].

⁸ Comparing Pedro Score categories (low < 5 vs. moderate 5–6 vs. high ≥ 7 score points; [14]) did not indicate significant BMD differences ($p > 0.416$) at LS-, FN-, and TH-ROI.

Table 3 Assessment of risk of bias for included studies

First author, year	Eligibility criteria	Random allocation	Allocation concealment	Inter group homogeneity	Blinding subjects	Blinding personnel	Blinding assessors	Participation \geq 85% allocation	Intention to treat analysis ^a	Between group comparison	Measure of variability	Total score
Bemben, 2000	Y	1	0	1	0	0	0	0	0	1	1	4
Bemben 2010	Y	0	0	1	0	0	0	1	1	1	1	5
Chilibeck, 2002	Y	1	1	1	0	0	0	0	1	1	1	6
de Oliveira, 2018	Y	1	1	1	0	0	1	1	1	1	1	8
Duff, 2016	Y	1	1	1	1	0	1	0	1	1	1	8
Hartard, 1996	Y	0	0	1	0	0	0	1	1	1	1	5
Kerr, 1996	Y	1	0	1	0	0	0	0	1	1	1	5
Kohrt, 1997	Y	0	0	1	0	0	0	0	1	1	1	4
Maddalozzo, 2007	Y	1	0	1	0	0	0	1	1	1	1	6
Nelson, 1994	Y	1	0	1	0	0	0	1	1	1	1	6
Nicholson, 2015	Y	1	1	1	0	0	0	1	1	1	1	7
Orsatti, 2013	Y	1	1	1	0	0	0	1	1	1	1	7
Pruitt, 1995	Y	1	0	1	0	0	0	0	1	1	1	5
Pruitt, 1992	Y	0	0	1	0	0	0	1	1	1	1	5
Rhodes, 2000	Y	1	0	1	0	0	0	1	1	1	1	6
Sinaki, 1989	Y	1	0	1	0	0	0	1	1	1	1	6
Woo, 2007	Y	1	1	1	0	0	1	1	1	1	1	8

^aThe point is awarded not only for intention to treat analysis but also when all subjects for whom outcome measures were available received the treatment or control condition as allocated

Random-effects Analysis of LS data

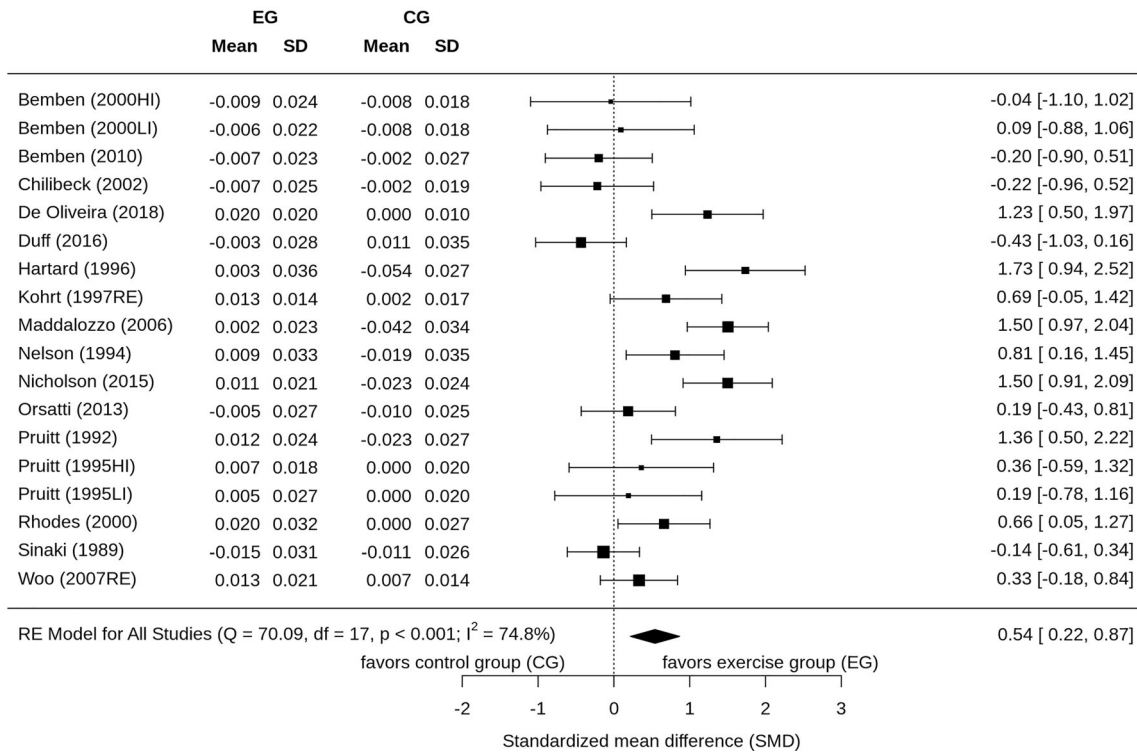


Fig. 2 Forest plot of meta-analysis results at the LS. The data are shown as pooled standard mean difference (SMD) with 95% CI for changes in exercise and control groups

3 a high (≥ 300 reps/session) exercise volume/session (Table 2). The mixed-effects analysis did not demonstrate any significant difference between the sub-groups ($p = 0.697$).

Effect of exercise on FN-BMD

Fifteen studies with 18 exercise groups evaluated the effect of exercise on FN-BMD. In summary, the exercise interventions resulted in significant ($p = 0.005$), but low effects sizes (SMD 0.22, 95% CI 0.07–0.38) (Fig. 4). There was a negligible level of heterogeneity in estimates of the exercise effect ($I^2 = 0.0\%$,

$Q = 13.0$). Sensitivity analysis of imputation determined that even in the worst case (i.e., imputing maximum SD), there is a significant effect (SMD 0.17, 95% CI 0.02–0.33, $p = 0.027$). Results listed in Fig. 4 are based on mean value imputation.

The funnel plot indicates evidence for a publication bias (Fig. 5). The regression ($p = 0.604$) and rank correlation test ($p = 0.601$) for funnel plot asymmetry indicate relevant asymmetry. The analysis indicates missing studies on the lower right-hand side. A trim and fill analysis resulted in slightly higher effects sizes (SMD 0.26, 95% CI 0.11–0.41), after adjusting for publication bias.

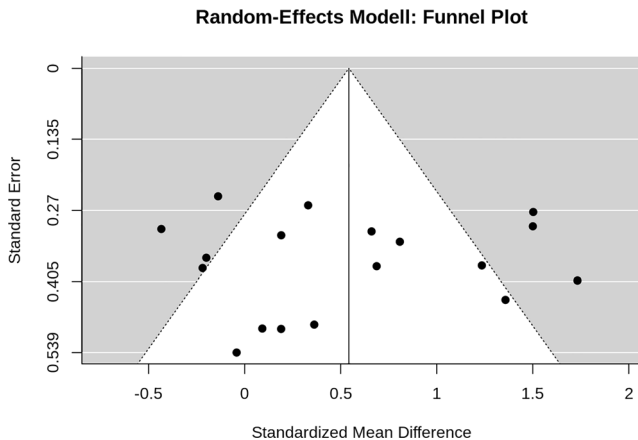


Fig. 3 Funnel plot of the DRT studies that address LS BMD

Sub-group analyses for FN-BMD

Intervention duration

Of 18 groups, 6 studies applied a short, 12 groups a moderate, and no group a long duration of the exercise intervention (Table 2). A mixed-effects analysis did not observe significant differences between the sub-groups ($p = 0.694$).

Type of exercise

Ten groups worked with resistance training devices, two with free weights or resistance bands group, and six groups conducted a mix of both types (Table 2). A mixed-effects analysis

Random-effects Analysis of FN data

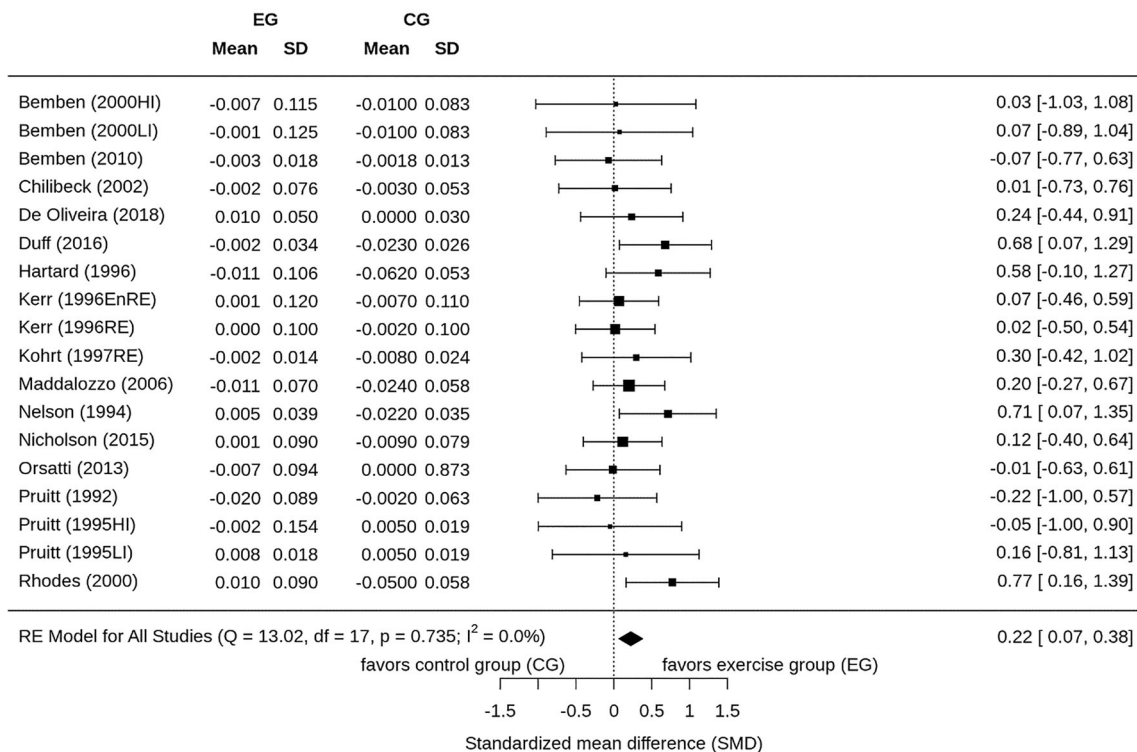


Fig. 4 Forest plot of meta-analysis results at the femoral neck. The data are shown as pooled standard mean difference (SMD) with 95% CI for changes in exercise and control groups

demonstrated no significant differences between the sub-groups ($p = 0.490$).

Training frequency

Five groups exercised fewer than 2 sessions per week and 13 groups exercised ≥ 2 sessions per week (Table 2). In contrast to LS-ROI, no significant difference was observed between the two groups ($p = 0.260$) from mixed-effects analysis.

Exercise intensity

Four groups exercised with low ($< 65\%$ 1RM), 7 with moderate ($65 < 80\%$ 1RM), and 7 with high relative intensity ($\geq 80\%$ 1RM) (Table 2). The mixed-effects analysis determined no significant differences between the sub-groups ($p = 0.279$).

Exercise volume/session (exercises \times sets \times reps)

Six groups applied a low (< 160 reps/session), 7 a moderate (160 to < 300 reps/session), and 5 a high (≥ 300 reps/session) exercise volume/session (Table 2). No significant differences between the sub-groups ($p = 0.373$) were demonstrated by the mixed-effects analysis.

Effect of exercise on TH-BMD

Nine studies with 11 exercise groups evaluated the effect of exercise on TH-BMD (Fig. 6). In summary, the pooled estimate of random effect analysis was 0.48, 95% CI 0.22–0.75. Level of heterogeneity between trials was low ($I^2 = 35.8\%$, $Q = 14.7$). Sensitivity analysis demonstrated that even in the worst case (i.e., maximum SD), there is a significant effect (SMD 0.44, 95% CI 0.22–0.65, $p = 0.001$). Results listed in Fig. 6 are based on mean value imputation.

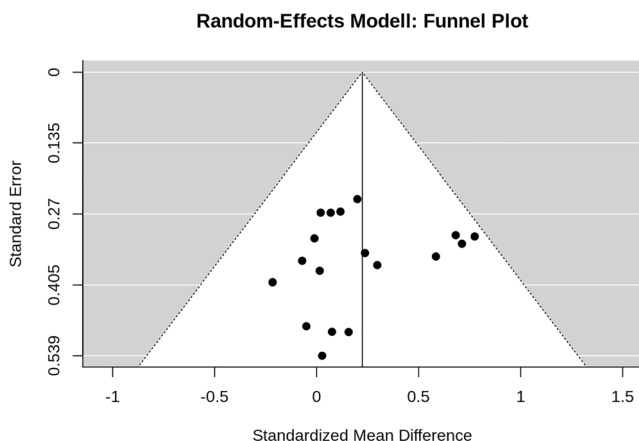


Fig. 5 Funnel plot of the DRT studies that address femoral neck BMD

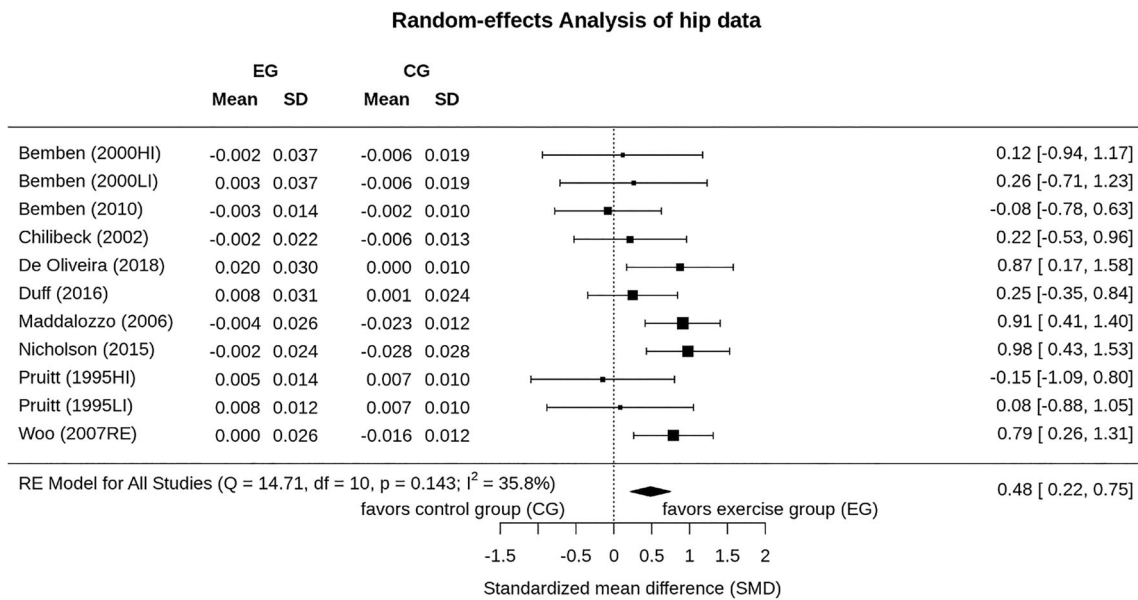


Fig. 6 Forest plot of meta-analysis results at the total hip. The data are shown as pooled standard mean difference (SMD) with 95% CI for changes in exercise and control groups

The funnel plot suggested positive evidence of publication bias (Fig. 7). The regression ($p = 0.013$), but not the rank correlation test ($p = 0.218$) for funnel plot asymmetry, demonstrated significant asymmetry. Comparable to FN-BMD, there is a lack of studies on the (lower) right-hand side. A trim and fill analysis resulted in considerably higher effects sizes after adjusting for publication bias (SMD 0.67, 95% CI 0.40–0.93).

Sub-group analyses for the TH-BMD

Intervention duration

Of 11 groups, 5 training groups were classified as short-term, 6 groups as moderate, and no training groups was categorized as long-term interventions (Table 2). A mixed-effects analysis

indicated no significant difference between the sub-groups ($p = 0.835$).

Type of exercise

Of 11 groups, 7 groups worked with resistance training devices, three with free weights or resistance bands group, and one group conducted a mix of both types (Table 2). A mixed-effects analysis indicated a significant difference between the sub-groups ($p = 0.013$). The sub-group analysis demonstrated the highest effects sizes (SMD 0.89, 95% CI 0.59–1.19) for the “free weight” training group (vs. RT devices 0.23, -0.09–0.55; vs. mix 0.25, -0.35–0.84).

Training frequency

Only two training groups applied low training frequency (< 2 sessions/week); nine training groups prescribed a high training frequency (≥ 2 sessions/week). Differences between the groups were significant ($p = 0.023$); sub-groups analysis provided higher effect sizes for the sub-group with lower training frequency (SMD 0.94, 95% CI 0.57–1.31 vs. high frequency 0.34, 0.07–0.61).

Exercise intensity

Four groups worked with low relative intensity (< 65% 1RM), four with moderate relative intensity (65–< 80% 1RM), and three with high relative intensity ($\geq 80\%$ 1RM) (Table 2). The mixed-effects analysis did not determine any significant differences between the sub-groups ($p = 0.090$).

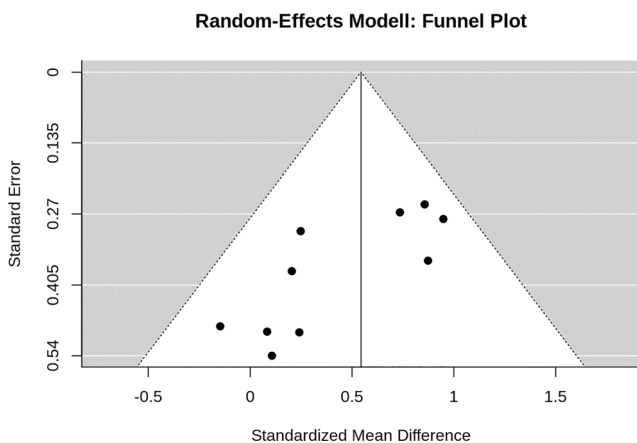


Fig. 7 Funnel plot of the DRT studies that address total hip BMD

Exercise volume/session (exercises × sets × reps)

Two groups exercised with low, six with moderate, and three with high exercise volume/session protocols (Table 2). No significant differences between the sub-groups ($p = 0.84$) were observed.

Discussion

In this systematic review and meta-analysis, we clearly confirmed the significant positive effects of DRT type exercise on BMD in postmenopausal women. However, effect sizes differed considerably between the regions of interest. While SMD for the LS-BMD (0.59) and total hip BMD (0.48) can be considered as moderate, the corresponding effect for the FN was quite low (SMD 0.22). We hypothesize that different loading configuration during DRT might predominately explain the different results at not only LS and FN but also FN and TH-ROI.⁹ One explanation for the significantly lower effect on the femoral neck region during resistance exercise might be the high stress level from everyday activities at this site. Due to leverage ratios, weight-bearing loads in one-legged stand situations such as walking result in high tensions of the abductor muscles and high stresses especially at the femoral neck region. Force measurements revealed corresponding loads of approximately three times the body mass (e.g., [42]). Also, RT-induced joint reaction forces might not have exceeded the threshold for bone adaptation, and hence, no exercise effect occurred.

Summarizing the few other meta-analyses [6, 8, 43, 44] that focus on the effect of RT on BMD at LS, or proximal femur ROIs, effect sizes vary considerably. Most of them reported negligible to low effects of RT on BMD, be it at the LS (SMD ≤ 0.24 [43]) or femoral neck (SMD ≤ 0.21 [9]). We observed higher effects sizes particularly for LS (SMD = 0.54) and TH-BMD (SMD 0.48). Due to the longer search period, we included more studies than most other systematic reviews but the main difference between the present study and previous analyses might be the more careful screening of eligibility [43] related to isolated DRT protocols. Of importance for the generation of exercise protocols, one analysis [6], which divided studies according to their exercise intensity, reported a missing effect for low “force” RT protocols (LS $n = 5$, MD -0.17% ; FN $n = 3$, MD -0.03% ; TH, $n = 3$: MD 0.21%). “High force” RT protocols ($\approx 60\%$ 1RM), however, demonstrated significant but low–moderate BMD effects at the LS ($n = 8$, MD 0.86%) and FN ($n = 8$, MD

1.03%) but not for BMD at TH ($n = 5$, MD 0.11% , 95% CI -0.06 – 0.29%). On the other hand, the study of Martyn-St-James et al. [8] that included only high intensity DRT studies and our sub-group analyses did not confirm the result that RT with moderate (65 – 80% 1RM) or high intensity ($\geq 80\%$ 1RM) is superior to RT protocols with low “force” [6].¹⁰ If anything, reviewing the data of the sub-analyses revealed some unexpected findings.

Intervention duration

Considering that the mechanical stress during RT induces changes of BMD in adults might be triggered predominately by remodeling, we hypothesized that studies below 9 months of length would not determine the full amount of new mineralized bone [45].¹¹ Nevertheless, addressing this issue by mixed-effects analysis, no significant BMD difference was observed between studies of longer and shorter durations, be it at the LS-, FN-, or TH-ROI. One may thus speculate that despite (Table 2) progressive increase of exercise intensity,¹² no relevant further exercise-induced effects occur after initial bone adaptation, be it by modeling or (fast) remodeling [24, 46]. We are unable to reject this conjuncture for pure RT protocols; however, previously published studies applying mixed exercise long-term protocols [47, 48] observed an ongoing effect of exercise on BMD at LS and FN up to 16 years.

Type of exercise

From a pragmatic point of view, it is important to determine whether specific RT devices are needed to generate successful exercise programs. In summary, in consistent favor for “free weight training” (vs. “device” and “mixed type” of training), we observed BMD differences at the LS-, FN-, and total hip-ROI, but albeit significant for the TH-ROI. This result is very welcome, not only due to the much lower material effort of free weight training. Of relevance for older people, free weights might be more favorable to increase function [49] and in particular leg extensor strength [50] with its crucial relevance on mobility limitations, disability, morbidity, and mortality [51–53].

TrFr

This parameter might be the most important aspect for designing an exercise protocol. Apart from its direct impact on the

⁹ Of importance, femoral neck ROI is a usually 15 mm slide across the middle-distal end of the femoral neck, while total hip ROI started slightly below the trochanter minor and include intertrochanter, trochanter, ward, and femoral neck ROI.

¹⁰ ...nor the results of higher effects on femoral neck ROI compared to total hip ROI.

¹¹ At least when bearing in mind that initial conditioning phases conducted by most studies might further shorten the period of over-threshold strain application...

¹² Apart from two studies, all the RT studies included progressively increased exercise intensity during the intervention (Table 2).

outcome addressed, training frequency (TrFr) immediately affects the feasibility of the program and thus participant compliance [54]. In summary, our result clearly indicates that lower net training frequency (< 2 sessions/week, s/w) demonstrated significantly higher effect sizes for BMD changes at the LS-ROI and TH-ROI¹³ versus higher training frequency (≥ 2 s/w). At the latest at this point, we have to subject our meta-analytical results to a critical review. Although some exercise studies did not detect significant BMD differences after exercising with varying TrFr¹⁴ [55–57], other studies [58–61] clearly demonstrated significantly higher effect sizes for BMD changes at LS and hip-ROI when applying TrFr of at least 2 s/w, compared with $1 < 2$ s/w. In these studies, lower TrFr was not only less favorable but showed no effects on BMD at LS or hip-ROIs at all. Nevertheless, there might be some explanations for at least similar BMD results after exercising with different TrFr. Firstly, one may speculate that higher intensity might compensate the effect of lower frequency or (vice versa) high frequency combined with high intensity might result in incomplete adaptation to exercise [10]. However, a sub-analysis combining training frequency with exercise intensity did not support this hypothesis. From the literature results mentioned above,¹⁵ it would also be conceivable that particularly during the early phase of an intervention, each bone-specific exercise protocol might trigger positive effects on BMD largely independent of TrFr. However, combining training frequency and intervention duration in another sub-analysis did not support this hypothesis, either. Finally, the rather low variance within net TrFr (≈ 1.5 – 3.5 s/w) might confound a proper result on this issue. However, summarizing the result on TrFr of the present study, from a sport-scientific point of view, it is hardly possible that in this reasonable range of TrFr, a lower TrFr triggers significantly higher effects on BMD changes compared with higher TrFr.

Exercise intensity

Another key parameter of exercise effects on a given outcome is “exercise intensity.” We categorized relative exercise intensity according to % 1RM (low $< 65\%$ vs. moderate $65 < 80\%$ vs. high intensity $\geq 80\%$ 1RM) as listed by the exercise trials. In summary, however, the sub-group analysis did not reveal significant differences between the groups. Although not consistently determined (e.g., [19, 57]), there is a high level of evidence [25, 62–64] that high exercise intensity is superior to

moderate or low exercise intensity for addressing BMD. The superiority of high intensity RT is strongly supported by basic research [65, 66], which indicates that the higher strain magnitude generated by higher deformation of the bone increases bone formation linearly to its deformation magnitude. Contrary to a fixed “bone adaptation threshold” at $1000 \mu\text{m}$, as suggested by the Mechanostat theory [67], other authors revealed that loading thresholds for modeling/remodeling vary between different skeletal sites, according to their habitual loading history [68–70]. However, as reported, we did not find any evidence for the superiority of high intensity RT at any ROI¹⁶ addressed by this study.

Exercise volume/session

Basically, there is a close interaction between exercise volume, in particular cycle number (i.e., number of reps) and strain magnitude (i.e., exercise intensity) [10]. With respect to bone physiology, the number of loading cycles is negligible when applying a high strain magnitude [71]; however, there is some evidence [5, 72] that higher cycle numbers might compensate for low to borderline strain magnitudes. In this context, Cullen et al. [73] demonstrated that 40 repetitions with a strain magnitude of $1000 \mu\text{m}$ did not relevantly affect bone formation rate, while 120 or 400 reps resulted in a significant increase in this parameter. However, addressing the relevance of exercise volume/session for BMD changes by our sub-analysis, we do not observe any relevant ($p \geq 0.373$) effect of this parameter.

Limitations

In summary, our evaluation of exercise characteristics with particular relevance on BMD to identify moderators of exercise effects on bone strength largely failed. Accordingly, we are unable to recommend any promising DRT protocols for bone strengthening. In his critical review, Gentil et al. [74] questioned the relevance and practical application of meta-analytic results in strength training. While this might be going too far, it is nevertheless obvious, however, that the complex interaction of exercise variables, training principles, and training conditions¹⁷ prevent, or at least aggravate, a proper analysis of single exercise parameters even when focusing on relative homogeneous types of exercise (i.e., DRT). This is even more the case when applying exercise regimens in the real world and not in laboratory-based, artificially supported study settings. In addition, the brief reporting on relevant

¹³ LS-BMD: 5 (low ExFr) vs. 13 (high ExFr) study groups TH: 2 vs. 9 study groups (see “Results”).

¹⁴ Ashe et al. [55]: 0–2 sessions/w.(s/w); Bailey et al. [56]: 0–7 s/w; Bemben et al. [57]: 2 vs. 3 s/w. However, studies did not adjust for subjects’ attendance; therefore, results should be interpreted with caution.

¹⁵ While exercise protocols of studies that did not detect difference for TrFr were 6, 8, and 12 months, the lengths of studies that detected significant higher BMD changes in favor of the higher TrFr were 1.5, 4, 12, and 16 years.

¹⁶ In detail, the subgroup analysis on this issue was inconsistent, with higher effects of low intensity for TH-BMD and moderate intensity for FN- and LS-BMD.

¹⁷ This might be the main difference to meta-analysis in the area of pharmaceutical agents or supplements (e.g. [75, 76], with their limited number of inherent modulators.

characteristics in research papers often does not provide sufficient information. Thus, meta-analyses might be an appropriate tool for determining the general effects of dedicated exercise on a given outcome, but their ability to distinguish between exercise parameters is more limited.

Some limitations and features of this work should be addressed to allow the reader to adequately interpret our findings and to follow our conclusions.

(1) We set out to determine the effect of preferably isolated DRT. DRT was defined as any kind of resistance exercise that involves joint movement and focuses on the development of musculoskeletal strength,¹⁸ correspondingly excluded studies with other types of exercise, be it as training components or (bone) specific warm-ups. However, in reality, our approach might not always be considered consistent. Indeed, we included a study that also applied short bouts of rowing [26]. While accepting that 10 min of intense rowing can no longer count as RT,¹⁹ we did not think that the joint reaction force character of rowing [26] would confound our results. In parallel, studies that applied cycling or stretching, i.e., exercises with no relevant mechanical impact on bone [79, 80], were included. (2) Although, with 17 studies, including 20 exercise and 18 control groups, our sample size of isolated DRT studies was higher than the sample size of recent meta-analyses, the statistical power might have been too low to address some dedicated issues by sub-group analyses. This limitation refers particularly to TH-ROI with considerably lower sample sizes. (3) There is a consistent lack of reporting of relevant DRT exercise parameters in the present literature; correspondingly, we are unable to evaluate all the promising exercise variables. Apart from absolute intensity (“effort”),²⁰ movement velocity²¹ was rarely reported [27]. However, strain rate (corresponding to movement velocity in DRT), as defined as alteration in strain magnitude per second during the acceleration or deceleration of loading ($\mu\Sigma/s$) is an important mechanical parameter. Turner et al. [81], for example, observed a linear increase in bone formation rate with higher strain rates when using a protocol with constant strain magnitude but varying strain rates. Von Stengel et al. [82] confirmed this finding for DRT, by comparing fast-explosive vs. slow movement velocity during high intensity DRT. The authors reported that BMD changes in the power training group (i.e., explosive concentric velocity) significantly exceeded the results of the resistance training group (i.e., TUT 4 s–0 s–4 s). Considering the importance of this parameter for bone strengthening and the easy

and safe applicability of high strain rates/fast movement velocity generated by joint reaction forces during DRT even in older, more vulnerable cohorts [83], more exercise studies should focus on this exercise variable. (4) Although we did our best to adequately classify our exercise characteristics according to exercise terminology or bone physiology, we admit that some of the categorizations (e.g., exercise volume/session) were made somewhat arbitrarily in order to ensure an appropriate distribution for comparisons. (5) Even after adjusting primary study outcomes (LS-, FN-, TH-BMD) for multiple testing, the significance of the results remained. However, following recent recommendations [84], we do not adjust on secondary outcomes (i.e., sub-analyses). (6) There is some evidence for a publication bias for LS-, FN-, and TH-BMD data. Due to the preference to report positive effects [85], the true effect size of exercise on BMD was in general considered lower for unadjusted data. However, the lack of studies in the (lower) right-hand corner of the funnel plot indicates that small-moderate size studies with positive effects are missing. Indeed, using trim and fill analysis [18], we determined no changes for the LS-BMD but higher effects sizes for FN and particular TH-BMD.

In conclusion, it is difficult to generate exercise recommendations on bone strengthening based on the meta-analytic results of the present exercise trials. Uncritical acceptance of the acquired meta-analytic data is certainly unwarranted in this context. Based on this experience, we conclude that dedicated, accurately designed randomized controlled exercise trials might be the more appropriate tool for addressing single exercise characteristics and thus generating exercise recommendations in the area of osteoporosis prevention and therapy.

Acknowledgments We would like to thank the Elsbeth Bonhoff Stiftung (Berlin, Germany), a non-profit organization for supporting the study. The present work was performed in (partial) fulfillment of the requirements for obtaining the degree “Dr. rer. biol. hum.” for the first author (Mahdieh Shojaa).

Funding information Open Access funding provided by Projekt DEAL. The study was funded by the Elsbeth Bonhoff Stiftung (Berlin, Germany).

Data availability The data that support the findings of this study are available from the corresponding author (WK), upon reasonable request.

Compliance with ethical standards

Conflicts of interest None.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative

¹⁸ In contrast to other researchers (e.g. [77]), we do solely include studies that apply progression (i.e. progressive resistance training as defined as “exercise in which a load is increased in predetermined steps” [78]).

¹⁹ By definition “strength” started at 30% 1RM [10]; relative intensity of 10 min of rowing (at least ≥ 250 reps) is below this cutoff.

²⁰ E.g., as defined according to the set endpoint approach of Steele et al. [38].

²¹ as reported in time under tension in different sections of the movement (e.g., concentric–isometric–eccentric).

Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

- Kemmler W, Bebenek M, Kohl M, Von Stengel S (2015) Exercise and fractures in postmenopausal women. Final results of the controlled Erlangen Fitness and Osteoporosis Prevention Study (EFOPS). *Osteoporos Int* 26:2491–2499
- Kemmler W, Haberle L, von Stengel S (2013) Effects of exercise on fracture reduction in older adults: a systematic review and meta-analysis. *Osteoporos Int* 24:1937–1950
- Beck BR, Daly RM, Singh MA, Taaffe DR (2016) Exercise and Sports Science Australia (ESSA) position statement on exercise prescription for the prevention and management of osteoporosis. *J Sci Med Sport* 20:438–445
- Daly RM, Dalla Via J, Duckham RL, Fraser SF, Helge EW (2019) Exercise for the prevention of osteoporosis in postmenopausal women: an evidence-based guide to the optimal prescription. *Braz J Phys Ther* 23:170–180
- Kemmler W, Stengel V (eds) (2019) The role of exercise on fracture reduction and bone strengthening. Avademic Press, London
- Howe TE, Shea B, Dawson LJ et al (2011) Exercise for preventing and treating osteoporosis in postmenopausal women. *Cochrane Database Syst Rev*:CD000333
- Kelley GA, Kelley KS, Kohrt WM (2012) Effects of ground and joint reaction force exercise on lumbar spine and femoral neck bone mineral density in postmenopausal women: a meta-analysis of randomized controlled trials. *BMC Musculoskelet Disord* 13:177
- Martyn-St. James M, Caroll S (2006) High intensity resistance training and postmenopausal bone loss: a meta-analysis. *Osteoporos Int* 17:1225–1240
- Zhao R, Zhao M, Xu Z (2015) The effects of differing resistance training modes on the preservation of bone mineral density in postmenopausal women: a meta-analysis. *Osteoporos Int* 26:1605–1618
- Weineck J (2019) *Optimales training*. Spitta-Verlag, Erlangen
- Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 151(264–269):W264
- de Morton NA (2009) The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Aust J Physiother* 55:129–133
- Sherrington C, Herbert RD, Maher CG, Moseley AM (2000) PEDro. A database of randomized trials and systematic reviews in physiotherapy. *Man Ther* 5:223–226
- Ribeiro de Avila V, Bento T, Gomes W, Leitao J, Fortuna de Sousa N (2018) Functional outcomes and quality of life after ankle fracture surgically treated: a systematic review. *J Sport Rehabil* 27:274–283
- Cochrane (2016) *Cochrane handbook for systematic reviews of interventions*. In: Higgins J, Green S (eds) *The Cochrane Collaboration*
- R_Development_Core_Team (2019) *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing
- Viechtbauer W (2010) Conducting meta-analyses in R with the metafor package. *J Stat Softw* 36:1–48
- Duval SJ, Tweedie RL (2000) A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *JASA* 95:89–98
- Bemben DA, Feters NL, Bemben MG, Nabavi N, Koh ET (2000) Musculoskeletal responses to high- and low-intensity resistance training in early postmenopausal women. *Med Sci Sports Exerc* 32:1949–1957
- Bemben DA, Palmer IJ, Bemben MG, Knehans AW (2010) Effects of combined whole-body vibration and resistance training on muscular strength and bone metabolism in postmenopausal women. *Bone* 47:650–656
- Chilibeck PD, Davison KS, Whiting SJ, Suzuki Y, Janzen CL, Peloso P (2002) The effect of strength training combined with bisphosphonate (etidronate) therapy on bone mineral, lean tissue, and fat mass in postmenopausal women. *Can J Physiol Pharmacol* 80:941–950
- de Oliveira LC, de Oliveira RG, de Almeida Pires-Oliveira DA (2019) Effects of whole-body vibration versus pilates exercise on bone mineral density in postmenopausal women: a randomized and controlled clinical trial. *J Geriatr Phys Ther* 42:E23–E31
- Duff WR, Kontulainen SA, Candow DG et al (2016) Effects of low-dose ibuprofen supplementation and resistance training on bone and muscle in postmenopausal women: a randomized controlled trial. *Bone Rep* 5:96–103
- Hartard M, Haber P, Ilieva D, Preisinger E, Huber JC (1996) Systematic strength training as a model of therapeutic intervention. *Arch Phys Med Rehabil* 75:21–28
- Kerr D, Morton A, Dick I, Prince R (1996) Exercise effects on bone mass in postmenopausal women are site-specific and load-dependent. *J Bone Miner Res* 11:218–225
- Kohrt WM, Ehsani AA, Brige SJ Jr (1997) Effects of exercise involving predominantly either joint-reaction or ground-reaction forces on bone mineral density in older women. *J Bone Miner Res* 12:1253–1261
- Maddalozzo GF, Widrick JJ, Cardinal BJ, Winters-Stone KM, Hoffman MA, Snow CM (2007) The effects of hormone replacement therapy and resistance training on spine bone mineral density in early postmenopausal women. *Bone* 40:1244–1251
- Nelson ME, Fiatarone MA, Morganti CM, Trice I, Greenberg RA, Evans WJ (1994) Effects of high-intensity strength training on multiple risk factors for osteoporotic fractures. A randomized controlled trial. *Jama* 272:1909–1914
- Nicholson VP, McKean MR, Slater GJ, Kerr A, Burkett BJ (2015) Low-load very high-repetition resistance training attenuates bone loss at the lumbar spine in active post-menopausal women. *Calcif Tissue Int* 96:490–499
- Orsatti F, Petri-Nahas E, Nahas-Neto J, Orsatti C, Teixeira A (2013) Effects of isoflavone and counter-resistance training on bone mineral density in postmenopausal women. *Rev Bras Cineantropom Desempenho Hum* 15:726–736
- Pruitt LA, Jackson RD, Bartels RL, Lehnhard HJ (1992) Weight-training effects on bone mineral density in early postmenopausal women. *J Bone Miner Res* 7:179–185
- Pruitt LA, Taaffe DR, Marcus R (1995) Effects of a one-year high-intensity versus low-intensity resistance training program on bone mineral density in older women. *J Bone Miner Res* 10:1788–1795
- Rhodes EC, Martin AD, Taunton JE, Donnelly M, Warren J, Elliot J (2000) Effects of one year of resistance training on the relation between muscular strength and bone density in elderly women. *Br J Sports Med* 34:18–22
- Sinaki M, Wahner HW, Offord KP, Hodgson SF (1989) Efficacy of nonloading exercises in prevention of vertebral bone loss in postmenopausal women: a controlled trial. *Mayo Clin Proc* 64:762–769
- Woo J, Hong A, Lau E, Lynn H (2007) A randomised controlled trial of Tai Chi and resistance exercise on bone health, muscle

- strength and balance in community-living elderly people. *Age Ageing* 36:262–268
36. Rhodes EC (1995) Effect of one year of resistance training on strength and bone density in elderly women. *Med Sci Sports Exerc* 27:S21
 37. Kerr D, Ackland T, Maslen B, Morton A, Prince R (2001) Resistance training over 2 years increases bone mass in calcium-replete postmenopausal women. *J Bone Miner Res* 16:175–181
 38. Steele J, Fisher J, Giessing J, Gentil P (2017) Clarity in reporting terminology and definitions of set end points in resistance training. *Muscle Nerve* 368–374:368–374
 39. Toigo M, Boutellier U (2006) New fundamental resistance exercise determinants of molecular and cellular muscle adaptations. *Eur J Appl Physiol* 97:643–663
 40. Hartard M, Haber P, Ilieva D, et al. (1994) Einfluß eines systematischen Krafttrainings auf die Knochendichte und auf das Niveau von Schmerzen und Befindlichkeit in einer Gruppe von Frauen mit Osteopenie. In Ließen H, Weiß M, Baum M (eds) *Regulations- und Repairmechanismen*. Köln
 41. Hoover DL, VanWye WR, Judge LW (2016) Periodization and physical therapy: bridging the gap between training and rehabilitation. *Phys Ther Sport* 18:1–20
 42. Bergmann G, Deuretzbacher G, Heller M, Graichen F, Rohlmann A, Strauss J, Duda GN (2001) Hip contact forces and gait patterns from routine activities. *J Biomech* 34:859–871
 43. Kelley GA, Kelley KS, Tran ZV (2001) Resistance training and bone mineral density in women: a meta-analysis of controlled trials. *Am J Phys Med Rehabil* 80:65–77
 44. Zhao R, Zhang M, Zhang Q (2017) The effectiveness of combined exercise interventions for preventing postmenopausal bone loss: a systematic review and meta-analysis. *J Orthop Sports Phys Ther* 47:241–251
 45. Eriksen EF (2010) Cellular mechanisms of bone remodeling. *Rev Endocr Metab Disord* 11:219–227
 46. Frost HM (1988) Vital biomechanics: proposed general concepts for skeletal adaptations to mechanical usage. *Calcif Tissue Int* 42:145–156
 47. Kemmler W, Engelke K, von Stengel S, Weineck J, Lauber D, Kalender WA (2007) Long-term four-year exercise has a positive effect on menopausal risk factors: the Erlangen Fitness Osteoporosis Prevention Study. *J Strength Cond Res* 21:232–239
 48. Kemmler W, Engelke K, von Stengel S (2016) Long-term exercise and bone mineral density changes in postmenopausal women—are there periods of reduced effectiveness? *J Bone Miner Res* 31:215–222
 49. Haff GG (2000) Roundtable discussion: machines versus free weights. *Strength Conditioning J* 22:18–30
 50. Schott N, Johnen B, Holfelder B (2019) Effects of free weights and machine training on muscular strength in high-functioning older adults. *Exp Gerontol* 122:15–24
 51. Newman AB, Kupelian V, Visser M, Simonsick EM, Goodpaster BH, Kritchevsky SB, Tylavsky FA, Rubin SM, Harris TB, on Behalf of the Health, Aging and Body Composition Study Investigators (2006) Strength, but not muscle mass, is associated with mortality in the health, aging and body composition study cohort. *J Gerontol A Biol Sci Med Sci* 61:72–77
 52. Roshanravan B, Patel KV, Fried LF, Robinson-Cohen C, de Boer IH, Harris T, Murphy RA, Satterfield S, Goodpaster BH, Shlipak M, Newman AB, Kestenbaum B, for the Health ABC study (2017) Association of muscle endurance, fatigability, and strength with functional limitation and mortality in the health aging and body composition study. *J Gerontol A Biol Sci Med Sci* 72:284–291
 53. Visser M, Goodpaster BH, Kritchevsky SB, Newman AB, Nevitt M, Rubin SM, Simonsick EM, Harris TB, for the Health ABC Study (2005) Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. *J Gerontol A Biol Sci Med Sci* 60:324–333
 54. Wagner P (2000) *Aussteigen oder Dabeibleiben?* Institute of Sports Sciences, Technical University Darmstadt, Darmstadt, p 298
 55. Ashe MC, Gorman E, Khan KM, Brasher PM, Cooper DM, McKay HA, Liu-Ambrose T (2013) Does frequency of resistance training affect tibial cortical bone density in older women? A randomized controlled trial. *Osteoporos Int* 24:623–632
 56. Bailey CA, Brooke-Wavell K (2010) Optimum frequency of exercise for bone health: randomised controlled trial of a high-impact unilateral intervention. *Bone* 46:1043–1049
 57. Bembem DA, Bembem MG (2011) Dose-response effect of 40 weeks of resistance training on bone mineral density in older adults. *Osteoporos Int* 22:179–186
 58. Cussler EC, Going SB, Houtkooper LB, Stanford VA, Blew RM, Flint-Wagner HG, Metcalfe LL, Choi JE, Lohman TG (2005) Exercise frequency and calcium intake predict 4-year bone changes in postmenopausal women. *Osteoporos Int* 16:2129–2141
 59. Kemmler W, von Stengel S (2013) Exercise frequency, health risk factors, and diseases of the elderly. *Arch Phys Med Rehabil* 94:2046–2053
 60. Kemmler W, von Stengel S (2014) Dose-response effect of exercise frequency on bone mineral density in post-menopausal, osteopenic women. *Scand J Med Sci Sports* 24:526–534
 61. Kemmler W, von Stengel S, Kohl M (2016) Exercise frequency and bone mineral density development in exercising postmenopausal osteopenic women. Is there a critical dose of exercise for affecting bone? Results of the Erlangen Fitness and Osteoporosis Prevention Study. *Bone* 89:1–6
 62. Maddalozzo GF, Snow CM (2000) High intensity resistance training: effects on bone in older men and women. *Calcif Tissue Int* 66:399–404
 63. Taaffe DR, Pruitt L, Pyka G, Guido D, Marcus R (1996) Comparative effects of high- and low-intensity resistance training on thigh muscle strength, fiber area, and tissue composition in elderly women. *Clin Physiol* 16:381–392
 64. Vincent KR, Braith RW (2002) Resistance exercise and bone turnover in elderly men and women. *Med Sci Sports Exerc* 34:17–23
 65. Hsieh YF, Turner CH (2001) Effects of loading frequency on mechanically induced bone formation. *J Bone Miner Res* 16:918–924
 66. Rubin CT, Lanyon LE (1985) Regulation of bone mass by mechanical strain magnitude. *Calcif Tissue Int* 37:411–417
 67. Frost HM (1987) Bone mass and the mechanostat. A proposal. *Anat Rec* 219:1–19
 68. Hsieh YF, Robling AG, Ambrosius WT, Burr DB, Turner CH (2001) Mechanical loading of diaphyseal bone in vivo: the strain threshold for an osteogenic response varies with location. *J Bone Miner Res* 16:2291–2297
 69. Mosley JR, Lanyon LE (1998) Strain rate as a controlling influence on adaptive modeling in response to dynamic loading of the ulna in growing male rats. *Bone* 23:313–318
 70. Turner CH (1998) Three rules for bone adaptation to mechanical stimuli. *Bone* 23:399–407
 71. Rubin CT, Lanyon LE (1984) Regulation of bone formation by applied dynamic loads. *J Bone Joint Surg Am* 66:397–402
 72. Kemmler W, von Stengel S (2011) Exercise and osteoporosis-related fractures: perspectives and recommendations of the sports and exercise scientist. *Phys Sportmed* 39:142–157
 73. Cullen DM, Smith RT, Akhter MP (2001) Bone-loading response varies with strain magnitude and cycle number. *J Appl Physiol* 91:1971–1976
 74. Gentil P, Arruda A, Souza D, Giessing J, Paoli A, Fisher J, Steele J (2017) Is there any practical application of meta-analytical results in strength training? *Front Physiol* 8:1
 75. Wang YK, Zhang YM, Qin SQ, Wang X, Ma T, Guo JB, Zhu C, Luo ZJ (2018) Effects of alendronate for treatment of

- glucocorticoid-induced osteoporosis: a meta-analysis of randomized controlled trials. *Medicine (Baltimore)* 97:e12691
76. Tai V, Leung W, Grey A, Reid IR, Bolland MJ (2015) Calcium intake and bone mineral density: systematic review and meta-analysis. *BMJ* 351:h4183
 77. Martyn-St James M, Carroll S (2006) Progressive high-intensity resistance training and bone mineral density changes among premenopausal women: evidence of discordant site-specific skeletal effects. *Sports Med* 36:683–704
 78. Kent M (2007) *The Oxford dictionary of sports science & medicine*. Oxford University Press, Oxford
 79. Chubak J, Ulrich CM, Tworoger SS et al (2006) Effect of exercise on bone mineral density and lean mass in postmenopausal women. *Med Sci Sports Exerc* 38:1236–1244
 80. Olmedillas H, Gonzalez-Aguero A, Moreno LA, Casajus JA, Vicente-Rodriguez G (2012) Cycling and bone health: a systematic review. *BMC Med* 10:168
 81. Turner CH, Owan I, Takano Y (1995) Mechanotransduction in bone: role of strain rate. *Am J Physiol Endocrinol Metab* 269: E438–E442
 82. von Stengel S, Kemmler W, Lauber D, Weineck J, Kalender WA, Engelke K (2005) Power training is more effective than strength training to maintain bone mineral density in postmenopausal women. *J Appl Physiol* 99:181–188
 83. Mangione KK, Miller AH, Naughton IV (2010) Cochrane review: improving physical function and performance with progressive resistance strength training in older adults. *Phys Ther* 90:1711–1715
 84. Li G, Taljaard M, Van den Heuvel ER et al (2017) An introduction to multiplicity issues in clinical trials: the what, why, when and how. *Int J Epidemiol* 46:746–755
 85. Sterne JA, Sutton AJ, Ioannidis JP et al (2011) Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 343:d4002

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.