

The effects of exercise on Bone Mineral Density in men: a protocol for a systematic review and meta-analysis of randomised controlled trials

Blair R. Hamilton^{1,2,3}, Katherine A. Staines^{3,4}, George A. Kelley⁵, Kristi S. Kelley⁵,
Wendy M. Kohrt⁶, Yannis Pitsiladis^{1,7}, and Fergus M. Guppy *^{3,4}

¹Collaborating Centre of Sports Medicine, University of Brighton, UK

²School of Sport and Service Management, University of Brighton, UK

³Centre for Stress and Age-related Disease, University of Brighton, UK

⁴School of Pharmacy and Biomolecular Sciences, University of Brighton, UK

⁵Department of Biostatistics, School of Public Health, West Virginia University, USA

⁶Department of Medicine, University of Colorado Anschutz Medical Campus, USA

⁷Department of Movement, Human and Health Sciences, University of Rome "Foro Italico," Rome, Italy

May 19, 2020

Abstract

Exercise is a cost-effective, widely available intervention that has been reported to help maintain optimal bone mineral density (BMD) in men, however, consideration of exercise modality is needed if the aim is to promote skeletal health. A previous meta-analysis of randomised controlled trials observed a moderate benefit on femoral neck (FN) but no benefit on lumbar spine (LS) BMD. However, since that analysis more randomised controlled trials (RCTs) have been published and updated methods of meta-analysis have been developed and therefore an updated systematic review and meta-analysis is required. **Methods and analysis:** RCTs of ≥ 24 weeks and published in English up to 01/05/20 will be retrieved by searching 3 electronic databases, cross referencing and expert review. The primary outcome measures will be changes in FN and LS BMD and lower limb BMD. Risk of bias for each study will be assessed using the Cochrane Risk of Bias instrument for RCTs, while the strength of evidence for each outcome will be assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) instrument. Standardised effect sizes will be calculated from each study and pooled using the inverse heterogeneity (IVhet) model. **Trial Registration number:** CRD42020180441.

*Corresponding Author: Fergus Guppy (F.Guppy@brighton.ac.uk)

All authors have read and approved this version of the manuscript for pre-print.

Authors FMG @fergusguppy can be contacted on twitter

Please cite as: Hamilton BR, Staines, KA, Kelley, GA, Kelley, KS, Kohrt WM, Pitsiladis, Y and Guppy FM. (2020). The Effects of Exercise on Bone Mineral Density in Men: a Protocol for a Systematic Review and Meta-Analysis *SportRxiv*. <https://doi.org/10.31236/osf.io/y8nue>

1 Introduction

Osteoporosis is a skeletal disorder characterized by low bone mineral density (BMD), compromised bone strength and architectural damage of the bone [1]. A reported 8.9 million osteoporotic fractures occur in Europe each year [1] and it is estimated to affect 200 million people worldwide [2]. While osteoporosis is four-fold more common in women than men, there is evidence that men have more osteoporosis-related complications [3, 4], with 30% experiencing a fracture [5]. Men who suffer a proximal femur fracture are also younger, less healthy and have a 2-3 fold higher mortality [6, 7, 8] and morbidity [9] than their female counterparts.

Exercise is a cost-effective, widely available intervention that has been reported to help maintain optimal BMD in men [10]; however, consideration of exercise modality is needed if the aim is to promote skeletal health. For example, swimming and cycling have been shown to have little or no effect on BMD [11, 12], with some studies finding that these forms of exercise actually reduce BMD at both the lumbar spine (LS) and proximal femur, often lower than inactive controls [13, 14, 15]. Weight-bearing, high-impact and strength training exercises have been recognised widely to be the gold standard for long-term skeletal health, with a 60% reduction in hip-fracture in men who were physically active compared to inactive men [16]. A previous meta-analysis of randomised controlled trials by Kelley et al. in 2013 [10] investigated the effects of exercise on BMD in men. The authors observed that there was moderate benefit of exercise on femoral neck (FN) BMD ($g = 0.583$ [0.031, 1.135]) but no benefit on LS BMD ($g = 0.190$ [-0.036, 0.416]). They concluded that there was insufficient evidence at that time to recommend ground- and/or joint-reaction force exercise for improving and/or maintaining FN and LS BMD in men and recommended that well-designed randomised controlled trials in men should be performed to formulate any final recommendations.

While the results reported by Kelley et al. are noteworthy, they were limited to only three randomised controlled trials published up to August of 2011 and lacked assessment of BMD using peripheral quantitative computed tomography (pQCT) [10]. However, since that time, additional randomised controlled trials have been published [17, 18] and more robust methods for the conduct and interpretation of meta-analytic results have been developed [19, 20, 21, 22]. Furthermore, to the best of the authors knowledge, no systematic review of previous systematic reviews with meta-analysis or original systematic review with meta-analysis has been conducted on the effects of exercise on BMD in men. Finally, using previously developed guidelines for when to update a systematic review, it was decided that an updated review on this topic was needed [23]. Thus, given 1) the deleterious consequences of low BMD in men, 2) the potential benefits of certain types of exercise on BMD in men [10], 3) the lack of recent meta-analytic work in this area, 4) the use of more robust methods for conducting meta-analytic research [19, 20, 21, 22], and 5) decision tree analysis of when to update a systematic review [23] our aim is to update the systematic review with meta-analysis by Kelley et al. [10], whereby we will examine the effects of exercise on BMD in men.

2 Materials and Methods

2.1 Study Eligibility Criteria

As this meta-analysis aims to update the Kelley et al. [10] meta-analysis, the same *a priori* inclusion criteria will be employed as shown in Table 1, with additional studies identified from 01/08/2011. Studies not meeting the criteria outlined in Table 1 will be excluded from the analysis. Studies will be limited to randomised controlled trials (RCTs) as this is the only way to ensure that confounders that are understood can be controlled, as well as to eliminate the overestimation that has been described in non-RCTs [24, 25].

As this meta-analysis is interested in the effects of exercise on skeletal health, studies with multiple interventions (e.g. diet combined with exercise) will only be included as long as there is a comparative control group (e.g. diet only). Resistance training studies will be limited to those studies that include lower body activities.

This meta-analysis will adhere to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) [26]. The protocol for this meta-analysis is registered on PROSPERO (trial registration number: CRD42020180441 [27]).

2.2 Data Sources

Studies published up to 01/05/2020 will be retrieved from three electronic sources (PubMed, Embase, SportDiscus). Keywords relevant to all searches will include “exercise”, “bone”, “randomised”. Based on PRISMA guidelines [26], an example search strategy will be submitted as supplementary material on completion. The first author (BH) will conduct all electronic database searches. In addition to electronic database searches, cross-referencing from retrieved studies will also be conducted.

2.3 Study Records and Selection

All studies will be imported into EndNote (EndNote X9.3.1, Clarivate Analytics, USA) and duplicates removed by the first author (BH). A copy of the database will then be provided to the last author (FMG) for duplicate screening. The first and last authors (BH and FMG) will select studies independent of each another. Multiple studies will be handled by including only the most recently published articles. The screeners will not be blinded to either the journal titles or to the study authors/affiliations. Reasons for exclusion will be coded based on one or more of the following; 1) inappropriate population, 2) inappropriate intervention, 3) inappropriate comparison(s), 4) inappropriate outcome(s), 5) inappropriate study design or 6) other. On completion, the screeners will meet to discuss their selections and reconcile any discrepancies by consensus. If agreement cannot be achieved, the second author (KS) will provide a recommendation. The agreement rate, prior to the reconciliation of any discrepancies, will be calculated using Cohen’s κ statistic [28]. The precision of searches will also be calculated as the number of studies included divided by the number of studies screened (less duplicates) [29]. We will then calculate the number-needed-to-screen (NNS) by taking the reciprocal of the precision [29]. A flow diagram of this process and a full list of studies will be made available as a supplementary file on submission of Stage 2.

2.4 Data Abstraction

Prior to data abstraction an electronic code book will be developed by the first (BH) and last (FMG) authors with input from all others. The data will be coded based on the following major categories; 1) study characteristics (e.g., author, journal, year, etc.), 2) participant characteristics (e.g. age, height, mass, etc.), 3) intervention details (e.g., type, length, frequency, etc.), and 4) outcome characteristics (e.g., sample sizes, baseline/post exercise means and SDs, etc.). The first (BH) and last (FMG) authors will extract all data independent of one another before meeting to resolve any discrepancies by consensus. When this cannot be achieved, the second author (KS) will provide a recommendation. Prior to this, the overall agreement rate will be assessed by Cohen’s κ statistic [28].

2.5 Outcome Measures

A priori primary outcome measures will include changes in FN and LS BMD measured by DXA or DPA and lower limb BMD measured by pQCT. Secondary, *a priori* outcomes will include changes in

body mass, body mass index, lean mass, and fat mass. Obtaining missing data will be attempted for all primary and secondary outcome measures if assessed by a study but the data provided proves inadequate to calculate an effect size. The corresponding author will be contacted three times via email with one-week between each communication. These communications will be tracked (e.g., dates, responses, success rates, etc.) to establish the success rate of this process.

2.6 Risk of Bias Assessment

Risk of bias for each study will be assessed using the recently revised Cochrane Risk of Bias instrument for randomised controlled trials (RoB 2) [30]. Using one or more signalling questions, the RoB 2 instrument assesses risk of bias in five distinct domains: (1) bias arising from the randomisation process, (2) bias due to deviations from intended interventions, (3) bias due to missing outcome data, (4) bias in measurement of the outcome, and (5) bias in selection of the reported result. Based on signalling questions, each domain is assessed as either ‘low risk’, ‘high risk’, or ‘some concerns’. Based on responses to each domain, the overall risk of bias for each study is then assessed as either ‘low risk’, ‘high risk’, or ‘some concerns’. We chose to use this risk of bias instrument over the various study quality instruments, including those focused on exercise intervention studies [31, 32] given the difficulty of the latter in differentiating between the quality of reporting and the quality in the conduct of a study [30]. No studies will be excluded from the analysis based on risk of bias assessment [33]. The first (BH) and last (FMG) authors will undertake the risk of bias assessment independently of one another, before meeting to resolve any discrepancies by consensus. Where this cannot be achieved, the second author (KS) will provide a recommendation.

2.7 Statistical Analysis

2.7.1 Calculation of effect sizes

The *a priori* primary outcomes for this meta-analysis will be changes in FN and LS BMD, calculated using the Hedges standardised mean difference effect size (ES), g , adjusted for small sample sizes [34]. The g for each group will be calculated as the change score difference (absolute or relative) in the exercise group minus the change score difference for the control group, divided by the pooled standard deviation. If this information is not available, g will be calculated using procedures described by Follmann et al. [35]. For studies reporting multiple post-intervention time points, we plan on calculating g based on baseline and the final time point closest to the end of the intervention.

2.7.2 Effect size pooling

Results will be pooled using the recently developed inverse heterogeneity (IVhet) model [19], a model which has been shown to be more robust than the traditional random-effects model employed by Kelley et al. [10]. Two-tailed z - α values ≤ 0.05 and non-overlapping 95% confidence intervals will be considered statistically significant.

2.7.3 Heterogeneity and Inconsistency

For each pooled outcome, heterogeneity will be assessed using Q [36], with an α level of ≤ 0.10 considered to represent statistically significant heterogeneity. Inconsistency will be assessed using I^2 , an extension of Q . For this meta-analysis, inconsistency will be categorised as very low ($\leq 25\%$), low (25-50%), moderate (50-75%) or large ($\geq 75\%$) [36]. Absolute between-study heterogeneity will be assessed using tau squared (τ^2). In addition, influence analysis will be conducted by removing each study from

our analysis once in order to examine the effect of that study on the overall findings. Furthermore, the recently developed *E-value* will be used as a form of sensitivity analysis against unknown and/or unmeasured confounders [21]. The E-value is defined as the “minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates” [21]. Larger E-values (≥ 2.0) suggest that a substantial amount of unmeasured confounding would be necessary to explain away a treatment effect [21]. Given the expected small sample size, no subgroup or meta-regression analysis will be conducted.

2.8 Meta-biases

Small-study effects (publication bias, etc.) will be assessed qualitatively using the Doi plot and quantitatively using the Luis Furuya-Kanamori index (LFK index) [22, 37]. The Doi plot has been suggested to be more intuitive than the funnel plot and the LFK index more robust than the commonly used Egger’s regression-intercept test [22, 37]. LFK values within ± 1 , greater than ± 1 but within ± 2 , and greater than ± 2 are considered to represent no, minor, and major asymmetry, respectively [22].

2.9 Strength of evidence

Strength of findings for each outcome will be assessed using the most recent version of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) for meta-analysis tool [38]. Quality of evidence was assessed across the domains of risk of bias, consistency, directness, precision and publication bias. Quality was judged as high (further research is very unlikely to change our confidence in the estimate of effect), moderate (further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate), low (further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate), or very low (very uncertain about the estimate of effect) [38].

2.10 Software used for analysis

All data will be analysed using Meta XL (version 5.3). Upon completion, all data will be made available as supplementary material.

3 Contributions

Conceptualization, FMG. and KS; methodology, BH, FMG and GK; writing—original draft preparation, BH and FMG; writing—review and editing, ALL

4 Funding Information

This research received no external funding

5 Abbreviations

The following abbreviations are used in this manuscript:

BMD	Bone Mineral Density
DPA	Dual-energy Photon Absorptiometry
DXA	Dual-energy X-ray Absorptiometry
ES	effect size
FN	Femoral Neck
g	Hedges standardised mean difference effect size
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IVhet	Inverse heterogeneity
LFK	Luis Furuya-Kanamori
LS	Lumbar Spine
NNS	Number-needed-to-screen
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
pQCT	peripheral Quantified Computer Tomography
RCT	Randomised Controlled Trial
SD	Standard Deviation

Ünlühızcı K. Osteoporosis: Unawareness or Ignorance? *Erciyes Med J.* **2019**; *41* (1):1-3. DOI: 10.14744/etd.2019.19005

Sözen T, Özişik L, Başaran NÇ. An overview and management of osteoporosis. *Eur J Rheumatol.* **2017**; *4*(1):46. DOI: 10.5152/eurjrheum.2016.048.

Center JR, Nguyen TV, Schneider D, Sambrook PN, Eisman JA. Mortality after all major types of osteoporotic fracture in men and women: an observational study. *Lancet.* **1999**; *353*(9156):878-82. DOI: 10.1016/S0140-6736(98)09075-8

Feldstein A, Elmer PJ, Orwoll E, Herson M, Hillier T. Bone mineral density measurement and treatment for osteoporosis in older individuals with fractures: a gap in evidence-based practice guideline implementation. *Arch Intern Med.* **2003**; *163*(18):2165-72. DOI: 10.1001/archinte.163.18.2165

Reginster J-Y, Burlet N. Osteoporosis: a still increasing prevalence. *Bone.* **2006**; *38*(2):S4-9. DOI: 10.1016/j.bone.2005.11.024

Forsén L, Sögaard A, Meyer H, Edna T-H, Kopjar B. Survival after hip fracture: short-and long-term excess mortality according to age and gender. *Osteoporos Int.* **1999**; *10*(1):73-8. DOI: 10.1007/s001980050197

Haentjens P, Magaziner J, Colón-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B, et al. Meta-analysis: excess mortality after hip fracture among older women and men. *Intern Med.* **2010**; *152*(6):380-90. DOI: 10.7326/0003-4819-152-6-201003160-00008.

Holt G, Smith R, Duncan K, Hutchison J, Gregori A. Gender differences in epidemiology and outcome after hip fracture: evidence from the Scottish Hip Fracture Audit. *J Bone Joint Surg Br* **2008**; *90*(4):480-3. DOI: 10.1302/0301-620X.90B4.20264.

Sterling RS. Gender and race/ethnicity differences in hip fracture incidence, morbidity, mortality, and function. *Orthop Relat Res.* **2011**; *469*(7):1913-8. DOI: 10.1007/s11999-010-1736-3.

Kelley GA, Kelley KS, Kohrt WM. Exercise and bone mineral density in men: a meta-analysis of randomized controlled trials. *Bone.* **2013**; *53*(1):103-11. DOI: 10.1016/j.bone.2012.11.031.

Khan K, McKay H, Kannus P, Wark J, Bailey D, Bennell K. Physical activity and bone health: *Human Kinetics*; **2001**.

- Barry DW, Kohrt WM. BMD decreases over the course of a year in competitive male cyclists. *J Bone Miner Res.* **2008**; *23*(4):484-91. DOI: 10.1359/jbmr.071203
- Fredericson M, Ngo J, Cobb K. Effects of ball sports on future risk of stress fracture in runners. *Clin J Sport Med.* **2005**; *15*(3):136-41. DOI: 10.1097/01.jsm.0000165489.68997.60
- Fredericson M, Chew K, Ngo J, Cleek T, Kiratli J, Cobb K. Regional bone mineral density in male athletes: a comparison of soccer players, runners and controls. *Br J Sports Med.* **2007**; *41*(10):664-8. DOI: 10.1136/bjsm.2006.030783
- Nagle KB, Brooks MA. A systematic review of bone health in cyclists. *Sports Health.* **2011**; *3*(3):235-43. DOI:10.1177/1941738111398857
- Kujala UM, Kaprio J, Kannus P, Sarna S, Koskenvuo M. Physical activity and osteoporotic hip fracture risk in men. *Arch Intern Med.* **2000**; *160*(5):705-8. DOI: 10.1001/archinte.160.5.705
- Allison SJ, Folland JP, Rennie WJ, Summers GD, Brooke-Wavell K. High impact exercise increased femoral neck bone mineral density in older men: a randomised unilateral intervention. *Bone.* **2013**; *53*(2):321-328. DOI: 10.1016/j.bone.2012.12.045
- Bolam KA, Skinner TL, Jenkins DG, Galvão DA, Taaffe DR. The Osteogenic Effect of Impact-Loading and Resistance Exercise on Bone Mineral Density in Middle-Aged and Older Men: A Pilot Study. *Gerontology.* **2015**; *62*(1):22-32. DOI: 10.1159/000435837
- Doi SA, Barendregt JJ, Khan S, Thalib L, Williams GM. Advances in the meta-analysis of heterogeneous clinical trials I: The inverse variance heterogeneity model. *Contemp Clin Trials* **2015**; *45*: 130-138. DOI: 10.1016/j.cct.2015.05.009
- Doi SAR, Furuya-Kanamori L, Thalib L, Barendregt JJ. Meta-analysis in evidence-based healthcare: a paradigm shift away from random effects is overdue. *Int J Evid Based Healthc* **2017**; *15* 152-160. DOI: 10.1097/XEB.0000000000000125
- VanderWeele TJ, Ding P. Sensitivity analysis in observational research: Introducing the E-value. *Ann Intern Med* **2017**; *167* (4): 268-274. DOI: 10.7326/M16-2607
- Furuya-Kanamori L, Barendregt JJ, Doi SAR. A new improved graphical and quantitative method for detecting bias in meta-analysis. *Int J Evid Based Healthc* **2018**; *16*: 195-203. DOI: 10.1097/XEB.0000000000000000
- Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, Beyene J, Chang S, Churchill R, Dearness K, Guyatt G, Lefebvre C, Liles B, Marshall R, Martinez Garcia L, Mavergames C, Nasser M, Qaseem A, Sampson M, Soares-Weiser K, Takwoingi Y, Thabane L, Trivella M, Tugwell P, Welsh E, Wilson, EC, Panel fo updating guidance for systematic reviews (PUGS). When and how to update systematic reviews: consensus and checklist. *BMJ* **2016**; *354*: i3507. DOI: 10.1136/bmj.i3507
- Sacks HS, Chalmers TC, Smith H. Randomized versus historical controls for clinical trials. *Am J Med* **1982**; *72*:233-40. DOI: 10.1016/0002-9343(82)90815-4
- Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* **1995**; *273*:408-12. DOI:10.1001/jama.1995.03520290060030
- Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Plos Med* **2009**; *21*;6(7):e1000097. DOI: 10.1371/journal.pmed.1000097.
- Hamilton BR, Staines KA, Kelley GA, Kelley KS, Kohrt WM, Pitsiladis Y, Guppy FM. The effects of exercise on bone mineral density in men: a systematic review and meta-analysis of randomised controlled trials. *PROSPERO* **2020** CRD42020180441. Available from bluehttps://www.crd.york.ac.uk/

- Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* **1968**; *70*: 213–20. DOI: 10.1037/h0026256
- Lee E, Dobbins M, DeCorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. *BMC Med Res Methodol* **2012**; *12*: 51. DOI: 10.1186/1471-2288-12-51.
- Sterne JAC, Savovic J, Page MJ, Elbers, RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* **2019**; *366*: 14898. DOI: 10.1136/bmj.l4898.
- Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* **2003**; *83*: 713-721. DOI: 10.1093/ptj/83.8.713
- Smart NA, Waldron M, Ismail H, et al. Validation of a new tool for the assessment of study quality and reporting in exercise training studies: TESTEX. *Int J Evid Based Healthc* **2015**; *13*: 9-18. DOI: 10.1097/XEB.0000000000000020
- Ahn S, Becker BJ. Incorporating quality scores in meta-analysis. *J Educ Behav Stat* **2011**; *36*: 555–85. DOI: 10.3102/1076998610393968
- Hedges LV, Olkin I. Statistical methods for meta-analysis. *San Diego, CA: Academic Press*; **1985**. ISBN 0080570658, 9780080570655
- Follmann D, Elliot P, Suh I, et al. Variance imputation for overviews of clinical trials with continuous response. *Clin Epidemiol* **1992**; *45*: 769–73.
- Higgins JPT, Thompson, SG, Deeks, JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* **2003**; *327*: 557-60. DOI: 10.1136/bmj.327.7414.557
- Furuya-Kanamori L, Xu C, Lin LF, et al. P value-driven methods were underpowered to detect publication bias: analysis of Cochrane review meta-analyses. *J Clin Epidemiol* **2020**; *118*: 86-92. DOI: 10.1016/0895-4356(92)90054-Q
- Schünemann H, Brozek J, Guyatt GH, Oxman A. GRADE handbook for grading quality of evidence and strength of recommendation [updated October 2013]. *The GRADE Working* **2013**.

Table 1: Study eligibility criteria.

Criteria

- 1) Randomised trials with comparative control group
- 2) Published in English
- 3) Men 18 years of age and older
- 4) Participants not taking part in regular exercise prior to the study enrolment
- 5) Ground- and/or joint-reaction force exercise intervention of at least 24 weeks duration
- 6) Included in Kelley et al. [10] OR published since 01 August 2011
- 7) Data available for changes in FN and/or LS BMD as assessed by dual-energy x-ray absorptiometry (DXA), dual-photon absorptiometry (DPA), or peripheral quantitative computed tomography (pQCT)