J Appl Physiol
88: 1730–1736, 2000.

Exercise and bone mineral density in men: a meta-analysis

GEORGE A. KELLEY,[1] KRISTI S. KELLEY,[1] AND ZUNG VU TRAN[2]
[1]*Department of Kinesiology and Physical Education, Northern Illinois University, DeKalb, Illinois 60115; and* [2]*Department of Preventive Medicine and Biometrics, University of Colorado Health Sciences Center, Denver, Colorado 80262*

Kelley, George A., Kristi S. Kelley, and Zung Vu Tran. Exercise and bone mineral density in men: a meta-analysis. *J Appl Physiol* 88: 1730–1736, 2000.—The purpose of this study was to use the meta-analytic approach to examine the effects of exercise on bone mineral density (BMD) in men. A total of 26 effect sizes (ES) representing 225 subjects from 8 studies met the criteria for inclusion. When BMD sites assessed were specific to the sites loaded during exercise, increases of ~2.6% (2.1% in the exercisers and −0.5% in the controls) were found. These results were statistically significant (ES = 0.213, 95% bootstrap confidence interval = 0.007–0.452). Statistically significant ES changes were found for older (>31 yr) but not younger (<31 yr) adults, with differences between groups statistically significant (P = 0.04). Statistically significant changes were also observed at the femur, lumbar, and os calcis sites. The results of this study suggest that site-specific exercise may help improve and maintain BMD at the femur, lumbar, and os calcis sites in older men. However, the biological importance of the small changes observed for most outcomes, quality of studies, and limited data pool prevent us from forming any firm conclusion regarding the use of exercise for maintaining and/or improving BMD in men. Clearly, a need exists for additional studies.

men; osteoporosis; systematic review; review; physical activity

OSTEOPOROSIS AND LOW BONE mass are major public health problems affecting ~23,000,000 women ≥50 yr of age in the United States (28). However, osteoporosis and low bone mass are also a major public health problem among men, affecting ~5,000,000 individuals ≥50 yr of age in the United States (28). Approximately 5% of white, Asian, Hispanic, and American-Indian men 50–79 yr of age have osteoporosis, whereas 3.5% of black men 50–79 yr of age have the disease (28). For those individuals ≥80 yr of age, these numbers increase to ~24% among white men, 17% among black men, and 5% among Asian, Hispanic, and American-Indian men (28). In addition, it has been estimated that osteoporosis-related fractures represent 3% of all Medicare costs and that the lifetime risk of an osteoporotic fracture for white men ≥50 yr of age is ~13% (33).

However, this lifetime risk may be an underestimate, inasmuch as a recent study in Australia found that the residual lifetime fracture risk in 60-yr-old men with average life expectancy was 29% (18).

Exercise has been recommended as a nonpharmacological approach for maximizing bone mineral density (BMD) during the younger years as well as improving bone density by increasing and/or preventing the loss of bone during the older years (35). Exercise may be especially appropriate, since it is a low-cost intervention that is available to most of the general public. However, training studies examining the effects of exercise on BMD in men have led to conflicting results (2, 4–6, 13, 23–25, 31, 36). For example, of the 10 studies previously cited, only 39% of the sites assessed were reported as statistically significant and positive compared with a control group. One of the possible reasons for the lack of statistically significant results may be the small sample size that comprised many of these trials. As a result, the ability to detect meaningful differences may have been compromised. Meta-analysis is a quantitative approach in which individual study findings addressing a common problem are statistically integrated and analyzed (8, 14, 16, 29). It has been shown to be more accurate than the vote-counting approach (15) and may be especially useful when the number of studies is small and/or the sample sizes within each study are small (29). Although several meta-analyses have been conducted on the effects of exercise on bone density in women (3, 19–21, 37), we are not aware of any meta-analytic work dealing with the effects of exercise on BMD in men. Given the health-care consequences of low BMD in men and the potential for exercise to improve BMD, a need exists to use a quantitative approach to examine the effects of exercise on BMD in this population. Thus the purpose of this study was to use the meta-analytic approach to examine the effects of exercise on BMD in men.

METHODS

Data Sources

Computerized literature searches of articles indexed between January 1966 and December 1998 were performed using MEDLINE, Current Contents, Sport Discus, and Dissertation Abstracts International databases. The following keywords were used alone or in various combinations for computer searches: bone, exercise, physical activity, men, males, physical fitness, fitness, and osteoporosis. The titles and abstracts of studies identified in the computerized searches were examined to exclude any that were clearly irrelevant. The full text of the remaining articles was retrieved, and each paper was read to determine whether it contained information on the topic of interest. Because computer searches have

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

been shown to yield fewer than two-thirds of relevant articles (9), the reference lists from original and review articles were also reviewed to identify any studies that had not been previously identified and appeared to contain information on the topic of interest. Hand searching of selected journals was also performed. Furthermore, three experts on exercise and BMD (Dr. Charlotte Sanborn, Dr. David Nichols, and Dr. Christine Snow) reviewed our reference list and coding sheet for thoroughness and completeness.

Study Selection

Inclusion criteria for this study were as follows: *1*) randomized or nonrandomized trials that included a comparative nonexercise control group, *2*) exercise as the only intervention, *3*) adult men (mean study age ≥18 yr) as subjects, *4*) journal articles, dissertations, and masters theses published in the English-language literature, *5*) studies published and indexed between January 1966 and December 1998, *6*) BMD (relative value of bone mineral per measured bone area) assessed, and *7*) training studies lasting ≥16 wk. Only information that met the above criteria was included in our analysis. Thus, for example, if BMD was also assessed in women, we did not include this information, since it did not meet our inclusion criteria. Because dissertations may eventually become full-length journal articles, we cross-referenced between the two to avoid duplication. We did not include abstracts and conference papers from national meetings because of the paucity of data provided as well as the inability to obtain complete data from the authors. Studies published in foreign language journals were also not included because of the potential error in the translation and interpretation of findings. Studies that met our inclusion criteria were also examined to ensure that the same subjects were not included in more than one study (11). For the two studies that met our inclusion criteria but did not provide appropriate information on changes in BMD (4, 23), personal contact was made with the authors to retrieve such information.

Data Extraction

Coding sheets that could hold 242 items were developed and utilized in this investigation. In addition, coding instructions that described how to code each item on the coding sheet were developed and utilized. To avoid inter- and intracoder bias, all data were independently extracted by two authors. The authors then met and reviewed every data point for accuracy and consistency. Disagreements were resolved by consensus. The major categories of variables coded included study characteristics, physical characteristics of subjects, and primary and secondary outcomes.

Statistical Analysis

Primary outcomes. The primary outcome in this study was changes in BMD. Because of the various ways in which the authors reported data on changes in BMD and because we also wanted to maximize the number of studies and outcomes that could be included in our analysis, we used the standardized difference approach as our effect size (ES) measure. This measure provides one with a statistic similar to a *z*-score. Each ES was calculated by subtracting the change outcome (percent or absolute) in the exercise group from the change outcome in the control group and then dividing this difference by the pooled standard deviation of the exercise and control groups (16). The ES was then corrected for small-sample bias (16). For those studies that did not report change outcome variances, these were estimated using previously developed methods (12). In general, an ES of 0.20 is considered a small effect, 0.50 a moderate effect, and 0.80 a large effect (7). An ES of 0.50, for example, means that the exercise group differed from the control group by 0.5 SD in favor of the exercise group. By use of a *z*-score table, this means that the exercise group would do better than ~69% of the control group. Although the first metric of choice should be the one that has the most meaning to the reader, in this case, the percent change difference, we were unable to use this metric because of insufficient reporting of data for calculating within-group percent change variances for each study. However, to enhance interpretation, we also calculated percent change differences for each study. Because of the small sample size in this study, especially for subgroup analyses, bootstrap resampling (5,000 iterations) was used to generate 95% bootstrap confidence intervals (BCI) around mean ES changes for BMD (10). The bootstrap technique is a computer-intensive, nonparametric method of estimating the reliability of the original sample estimate, in this case, ES changes in BMD. By randomly drawing from the available sample, with replacement, samples the same size as the original are generated. Each time an observation is selected for a new sample, each of the elements of the original sample has an equal chance of being selected. This is similar to replicating each member of a sample 5,000 times (iterations). The main advantage of this approach is that the estimate desired is not based on some theoretical distribution but, rather, on the sample itself. This approach frees one from the constraints of the central limit theorem. The number of iterations chosen was based on previous research demonstrating that improvement of estimation accuracy was limited beyond 5,000 iterations (38). If the 95% confidence interval included zero (0.00), it was concluded that there was no statistically significant effect of exercise on BMD.

Heterogeneity of ES changes in BMD was examined using the Q statistic (16). A random-effects model was used if changes were significantly heterogeneous ($P \leq 0.05$), whereas a fixed-effects model was used in the absence of significant heterogeneity (26).

For studies that included multiple outcomes because of more than one group, net changes were initially treated as independent data points. However, to examine the influence (sensitivity) of each study on the overall results, analyses were performed with each study deleted from the model.

Publication bias (the tendency for journals and/or authors to publish studies that yield statistically significant results) was examined using Kendall's τ statistic (1). A statistically significant result ($P \leq 0.05$) was considered to be suggestive of publication bias. In addition, we used a semiquantitative approach (funnel plot) to examine potential publication bias (22). This was accomplished by plotting the sample size on the vertical axis and changes in blood pressure on the horizontal axis. Usually, smaller studies will be more dispersed at the bottom of the funnel, whereas larger studies will be more congregated at the top. A gap at the bottom of the funnel on the left side indicates that small studies yielding null or negative results may be missing.

Study quality was assessed using a three-item questionnaire designed to assess bias, specifically, randomization, blinding, and withdrawals/dropouts (17). The number of points possible ranged from a low of 0 to a high of 5. All questions were designed to elicit yes (1 point) or no (0 points) responses. The questionnaire took <10 min/study. The questionnaire has been shown to be valid (face validity) and reliable (researcher-interrater agreement, $r = 0.77$, 95% confidence interval = 0.60–0.86) (17).

Subgroup analyses. Subgroup analyses for primary outcomes were performed using ANOVA-like procedures for

meta-analysis (16). These procedures provide statistics for within (Q_w)- and between (Q_b)-group differences. If statistically significant within-group (Q_w) heterogeneity existed ($P < 0.05$), a random-effects model was used. If no statistically significant within-group (Q_w) heterogeneity existed, a fixed-effects model was used. ES changes in BMD were initially examined when the data were partitioned according to whether the BMD sites assessed were specific to the sites loaded during the exercise protocol. Subgroup analysis was also performed when the data were partitioned according to the age of the subjects (≤31 vs. >31 yr of age), limited to those results in which the BMD sites assessed were specific to the sites loaded. We chose 31 yr of age as our cut point because it has been reported that peak bone mass is attained as late as ~30 yr of age (32). Further analyses were performed on site-specific results for older subjects in relation to specific assessment location (femur, lumbar, os calcis) and study design (randomized controlled trial vs. nonrandomized controlled trial). Bootstrap resampling (5,000 iterations) (10) was used to generate 95% confidence intervals around ES changes for all subgroups. Randomization tests (5,000 iterations) were used to generate probability values for between-group differences (30). Randomization tests using 5,000 iterations can detect a probability as low as 0.002 (30).

Regression analysis. Potential associations between ES changes in BMD and initial BMD and length of training were examined using regression procedures previously described by Hedges and Olkin (16).

Secondary outcomes. Secondary outcomes (changes in body weight, body mass index, percent body fat, lean body mass, maximum oxygen consumption, resting heart rate, calcium intake) were calculated as the difference (exercise minus control) of the changes (initial minus final) in these mean values. The original metric was used for all secondary outcomes. For those studies in which variance estimation was necessary, these were accomplished using the same procedures used to estimate variances for BMD (12). Fixed- and random-effects models were used following the same procedures described for BMD.

Unless otherwise noted, values are means ± SD. The α-level for statistical significance was set at $P \leq 0.05$. Bonferroni adjustments were not made because of the increased risk of a type II error.

RESULTS

Study Characteristics

A total of 3,141 titles and abstracts were reviewed. From those, 26 ES representing a total of 225 subjects (135 exercise, 90 control) and 18 groups (10 exercise, 8 control) from 8 studies met the criteria for inclusion (2, 5, 6, 13, 24, 25, 31, 36). The per person time to code each study once ranged from 0.67 to 2.63 h (1.04 ± 0.66 h). Two studies were not included because we were unable to obtain data necessary for the calculation of an ES (4, 23). Thus our percent loss that met our inclusion criteria was 20. One study (34) was not included because it contained some of the same subjects from another study that met our inclusion criteria (24). We chose to include the latter study because more complete data were available for extraction. A general description of the studies is shown in Table 1. Four studies were conducted in the United States (5, 24, 25, 36), two in Australia (2, 31), and one each in Japan (13) and the United Kingdom (6). For the five studies that reported such information (2, 6, 24, 31, 36), percent dropout, defined as the number of subjects who did not complete the study, ranged from 0 to 31% in the exercise groups (10 ± 12%) and from 0 to 22% in the control groups (5 ± 10%). The number of subjects in which pre- and post-BMD measures were assessed and included in our analysis ranged from 3 to 28 in the exercise groups (14 ± 8) and from 7 to 21 in the control groups (11 ± 6). Study quality ranged from 0 to 3 (1 ± 1). Compliance, defined as the percentage of exercise sessions attended, was >90% for the two studies that reported this information (13, 24). Reliability for BMD assessment (coefficient of variation) ranged from 0.4 to ~5.0% (1.5 ± 1.0%).

Physical Characteristics of Subjects

A description of the physical characteristics for the exercise and control groups may be found in Table 2. Only one study reported that calcium supplementation was given to subjects (13). For the four studies that reported such information, three reported that none of the subjects was taking any type of pharmacological intervention other than calcium that could affect BMD (2, 6, 13), and one (5) reported that subjects were taking drugs that could affect BMD. For the two studies that reported information on cigarette smoking, one reported that none of the subjects smoked cigarettes (24), and another reported that some of the subjects smoked (13). Three studies reported that subjects had not been physically active before participation in the study (5, 13, 24), whereas it appeared that another study included subjects that had been previously active (2). The two studies that reported information on previous fractures reported that subjects had not had previous fractures at the site(s) assessed before participation in the study (24, 25). None of the studies reported information on the alcohol intake of the subjects.

Primary Outcomes

Individual ES changes for primary outcomes (changes in BMD) are shown in Table 3. Approximately 39% of the 26 ES were reported as statistically significant and positive by the authors. Initial BMD values for all sites assessed ranged from 0.700 to 1.400 g/cm^2 in the exercise groups (1.120 ± 0.188 g/cm^2) and from 0.670 to 1.290 g/cm^2 in the control groups (1.078 ± 0.168 g/cm^2). With use of a random-effects model because of statistically significant heterogeneity ($Q = 45.01$, $P = 0.008$), overall ES changes were not statistically significant (ES = 0.028, 95% BCI = −0.166 to 0.230). ES changes were equivalent to an exercise-minus-control improvement of 2% (1.6% increase in the exercising subjects and 0.4% decrease in the controls). With each study deleted from the model once, ES changes ranged from a low of 0.028 (95% BCI = −0.166 to 0.230) to a high of 0.199 (95% BCI = −0.131 to 0.489). Although there was no quantitative evidence supporting publication bias ($r = -0.23$, $P = 0.10$), funnel plot analysis was suggestive of this potential form of bias.

Table 1. *General description of included studies*

Study	Design/Subjects	Intervention	BMD Assessment
Bennell et al. (2)	CT that included 2 different male exercise groups (27 power athletes, 31 endurance athletes) and 27 male controls 17–26 yr of age	12 mo of participation in power (sprinting, jumping, multievents) or endurance (middle- or long-distance) sports	DEXA (Hologic QDR 1000W) at upper limb, lumbar spine (L_1–L_4), femur, and tibia/fibula
Braith et al. (5)	RCT consisting of 16 male heart transplant patients assigned to exercise ($n = 8$, 56 ± 6 yr) or control ($n = 8$, 52 ± 10 yr) group	6 mo of lumbar extension exercises 1 day/wk and 2 days/wk of upper and lower body resistance training consisting of 1 set of 10–15 repetitions at 50% of 1 RM for each exercise (10 exercises total)	DEXA (Lunar) at femoral neck, lumbar spine (L_2–L_3), and total body
Cohen et al. (6)	CT consisting of 17 male novice college oarsmen (age 19.5 ± 2.4 yr) and 8 age- and gender-matched controls (19.3 ± 1.6 yr)	7-mo training program consisting of rowing (8 h/wk), weight training (1 h/wk), and running (1 h/wk)	DEXA (Lunar DPX) at lumbar spine (L_1–L_4), femoral neck, greater trochanter, and Ward's triangle
Fujimura et al. (13)	CT consisting of 15 previously sedentary male subjects (23–31 yr) assigned to an exercise ($n = 8$, 24.6 ± 2.83 yr) or control ($n = 7$, 26.4 ± 3.17 yr) group	4 mo of resistance training performed 3 times/wk and consisting of 2–3 sets of 10 repetitions at 60–80% of 1 RM for 7–8 exercises/session	DEXA (DCS-3000) at lumbar spine, femoral neck, midradius, and total body
Menkes et al. (24)	CT consisting of 18 previously sedentary male subjects (50–70 yr) assigned to exercise ($n = 11$, 59 ± 7.2 yr) or control ($n = 7$, 55 ± 2.65 yr) group	4 mo of resistance training performed 3 times/wk and consisting of 1 set of upper body and 2 sets of lower body exercises performed for 15 repetitions for each exercise	DEXA (Lunar DPX) at lumbar spine (L_2–L_4), femoral neck, and total body
Michel et al. (25)	CT that included 10 male runners (65.3 ± 7.91 yr) and 10 age- and gender-matched controls (64.5 ± 7.27 yr)	5 yr of running partitioned by groups that decreased their mileage by >20% vs. <20% over 5-yr period	QCT (GE9800) at lumbar spine
Pritchard et al. (31)	RCT that included 40 overweight male subjects assigned to exercise ($n = 21$, 44.9 ± 6.5 yr) or control ($n = 19$, 43.3 ± 4.5 yr) group	1 yr of aerobic exercise of choice (walking, jogging, swimming, stationary cycling) performed ≥3 times/wk for 30 min at 65–75% of maximum heart rate	DEXA (Hologic QDR 1000W) of total body
Williams et al. (36)	CT consisting of 20 exercising men (38–68 yr, mean 49.2) and 10 male controls (41–58 yr, mean 47.0)	9 mo of a marathon training program partitioned into groups that ran ~21.9 and 10 miles/wk, respectively	SPA at right central os calcis

Values are means ± SD. BMD, bone mineral density; DEXA, dual-energy X-ray absorptiometry; SPA, single-photon absorptiometry; QCT, quantitative computed tomography; CT, controlled trial; RCT, randomized controlled trial; RM, repetition maximum.

Subgroup Analyses

When data were partitioned according to whether the BMD sites assessed were specific to the sites loaded during the exercise protocol, statistically significant within-group changes were found when the sites assessed were specific to the site loaded (ES = 0.213, 95% BCI = 0.007–0.452) but not when the sites assessed were not specific to the sites loaded (ES = −0.205, 95% BCI = −0.613 to 0.219). ES changes in BMD were equivalent to exercise-minus-control increases of ~2.6% (2.1% in the exercisers and −0.5% in the controls) when the sites assessed were specific to the sites loaded and ~0.3% (−0.06% in the exercisers and −0.3% in the controls) when the sites assessed were not specific to the sites loaded. Although it was not statistically significant, there was a trend for between-group changes to be greater when the sites assessed were specific to the sites loaded ($Q_b = 3.02$, $P = 0.10$). Further subgroup analyses were then performed by limiting analyses to only those sites in which the assessment of BMD was specific to the sites loaded. These are shown in Table 4. Within-group analysis demonstrated that statistically significant increases were found for older but not younger adults and that the differences between groups were statistically significant. These changes were equivalent to an exercise-minus-control improvement of ~6.7% for older subjects (4.2% increase in exercise groups and 2.5% decrease in controls) and 0.4% increase in younger subjects (1% increase in exercisers and 0.6% increase in controls). When the one study that

Table 2. *Initial physical characteristics*

	Exercise		Control	
Variable	*n*	Mean ± SD	*n*	Mean ± SD
Age, yr	10	40.8 ± 17.9	8	41 ± 17.0
Height, cm	9	176.2 ± 4.2	7	174.6 ± 3.7
Weight, kg	9	76.0 ± 6.4	7	78.1 ± 9.2
BMI, kg/m^2	9	24.6 ± 2.6	7	25.7 ± 3.4
Fat, %	5	17.0 ± 7.6	4	21.1 ± 6.7
LBM, kg	5	60.6 ± 4.7	4	58.9 ± 5.25
Calcium, mg/day	4	1,090 ± 259	3	892 ± 95

n, Number of groups reporting data; BMI, body mass index; LBM, lean body mass.

Table 3. *Individual ES from studies*

Study	Sites Assessed	No. Assessed	ES(d)	Var(d)
Bennell et al. (2)	Upper limb (power athletes)	42	−0.185	0.096
	Lumbar spine L_1–L_4 (power athletes)	42	0.844	0.107
	Femur (power athletes)	42	0.042	0.095
	Tibia/fibula (power athletes)	42	−0.222	0.096
	Upper limb (endurance athletes)	49	−0.199	0.084
	Lumbar spine L_1–L_4 (endurance athletes)	49	0.344	0.085
	Femur (endurance athletes)	49	−0.343	0.085
	Tibia/fibula (endurance athletes)	49	0.034	0.083
Braith et al. (5)	Total body	16	0.877	0.282
	Femoral neck	16	0.705	0.271
	Lumbar spine L_2–L_3	16	1.569	0.353
Cohen et al. (6)	Lumbar spine L_1–L_4	25	0.879	0.207
	Femoral neck	25	−0.257	0.186
	Greater trochanter	25	−0.568	0.194
	Ward's triangle	25	−0.265	0.186
Fujimura et al. (13)	Total body	15	−0.025	0.268
	Lumbar spine	15	−0.141	0.269
	Femoral neck	15	−0.203	0.270
	Midradius	15	0.807	0.297
Menkes et al. (24)	Total body	16	−0.126	0.255
	Lumbar spine L_2–L_4	16	0.099	0.254
	Femoral neck	16	0.270	0.257
Michel et al. (25)	Lumbar spine L_1 (<20% decrease)	13	0.858	0.465
Pritchard et al. (31)	Total body	40	−1.010	0.118
Williams et al. (36)	Os calcis (22 miles/wk)	17	1.048	0.284
	Os calcis (10 miles/wk)	23	0.260	0.179

No. assessed, sum of exercise and control subjects in which BMD was assessed; ES(d), effect size (ES) corrected for small sample bias; Var(d), variance of ES(d); <20%, <20% decrease in running mileage over course of study.

yielded two ES from heart transplant patients in the older group was deleted from the model, statistically significant within-group effects were again observed for older (ES = 0.442, 95% BCI = 0.207–0.799) but not younger (ES = 0.066, 95% BCI = −0.158 to 0.333) subjects. However, no statistically significant differences were observed between groups (Q_b = 2.26, P = 0.23). ES changes were equivalent to exercise-minus-control improvements of ~4.0% in older subjects (2% increase in exercisers and 2% decrease in controls). With ES results limited to only older adults, statistically significant within-group ES were found at the femur, lumbar, and os calcis sites, with no statistically significant between-group differences observed. ES changes in BMD were equivalent to exercise-minus-control improvements of ~5.9% at the femur site (4% increase in exercisers and 1.9% decrease in controls), 10.7% at the lumbar site (5.8% increase in exercisers and 4.9% decrease in controls), and 1.6% at the os calcis site (2.1% increase in exercisers and 0.5% increase in controls). With the one study in heart transplant patients deleted, there was an exercise-minus-control improvement of ~6.1% in lumbar BMD (1.0% in exercisers and −5.1% controls). Statistically significant within-group ES increases in BMD were also found when data were partitioned by study design. This was equivalent to exercise-minus-control improvements of ~13.5% for randomized trials (9.8% increase in exercisers and 3.7% decrease in controls) and 4.2% for nonrandomized trials (2% increase in exercisers and 2.2% decrease in controls). However, the percent changes found for randomized controlled trials were derived from the one study in heart transplant patients. No statistically significant ES differences were observed between groups.

Regression Analysis

No statistically significant associations were found between ES changes in BMD and length of training or initial BMD.

Secondary Outcomes

No statistically significant differences were found for any of the secondary outcomes. These included body weight (−0.3 kg, 95% BCI = −2.2 to 0.79 kg), body mass index (−0.9 kg/m^2, 95% BCI = −0.9 to 0.1 kg/m^2), and lean body mass (0.3 kg, 95% BCI = −0.2 to 0.6). Insufficient data were provided to assess changes in percent fat, maximum oxygen consumption, resting heart rate, and calcium intake.

Table 4. *Subgroup analyses*

Variable	No. of Studies	No. Assessed	No. of ES	ES(d)	95% BCI	Q_b
Age + site specific						
Older (<31 yr)	4	85	7	0.605	0.324 to 1.032	5.89 (0.04)*
Younger (≤31 yr)	4	131	13	0.066	−0.157 to 0.312	
Location + site specific (older only)						
Femur	2	32	2	0.482	0.270 to 0.705	0.32 (0.86)
Lumbar	3	45	3	0.749	0.099 to 1.327	
Os calcis	1	40	2	0.565	0.260 to 1.048	
Design + site specific (older only)						
RCT	2	56	2	1.082	0.705 to 1.569	1.98 (0.24)
CT	3	69	5	0.442	0.204 to 0.799	

No. assessed, sum of exercise and control subjects in which BMD was assessed; BCI, bootstrap confidence interval; Q_b, differences between groups, with P values in parentheses. *Statistically significant, $P < 0.05$.

DISCUSSION

The results of this study suggest that site-specific exercise may help improve and maintain BMD in older men. These results also support the notion that changes in BMD are specific to the sites loaded during exercise in older men. However, these results should be interpreted with caution, inasmuch as they were based on a very limited data pool. Perhaps little change was seen in BMD among younger subjects because subjects already possessed optimal levels of BMD and/or the loss of BMD generally occurs during the later, rather than during the younger, years. In addition, it may also be that more subjects in the older group had been sedentary for a longer period or were not able to ambulate very much. This may be especially true for the one study that included heart transplant patients, because when we deleted this study from our subgroup analysis the increase in BMD in the exercise group decreased from ~6% to 1% (5).

The fact that exercise is an inexpensive, nonpharmacological approach that is available to most of the general public makes this form of treatment appealing, especially given the other physiological and psychological benefits that may be derived from participation in exercise. However, it is important to realize that ~60% of adults in the United States do not regularly participate in adequate amounts of physical activity (27). Furthermore, only 16% of the US population between the ages of 18 and 64 yr report that they regularly participate in progressive-resistance exercise (27). In addition, one must also consider the potential adverse effects of exercise, e.g., arthritis, injury, and cardiac events. Unfortunately, we were unable to determine any potential adverse effects of exercise in this meta-analysis because only one study, limited to heart transplant patients, reported such information (5). The authors concluded that the exercise intervention was safe because it was not associated with an increase in rejection (5). Thus, although participation in a regular exercise program may be efficacious for improving and/or maintaining BMD in men, it may not be very effective in the "real world." Consequently, alternative nonpharmacological and pharmacological interventions for increasing and/or maintaining BMD may be necessary.

Despite the positive results observed in this study, the biological importance of these small changes might not be sufficient to recommend exercise alone as a nonpharmacological intervention. Because the primary reason for improving and/or maintaining BMD is to reduce fracture risk, one would like to know how much of an increase or prevention of loss in BMD is necessary to reduce the incidence of fracture. A recent longitudinal study has shown that femoral neck BMD in men was 24% lower in those with hip fractures and 12% lower in those with fractures of the vertebrae and upper limb (28a). The study also found that a decrease of 1 SD in femoral neck BMD was associated with a 2.3, 1.9, and 1.5 increase in the odds of fracture at the hip, vertebrae, and upper limbs, respectively (28a). The smaller results observed in our study suggest that if exercise is recommended, it should be done only in conjunction with other types of nonpharamcological and/or pharmacological interventions. Furthermore, we are not aware of any consensus as to how exercise-induced increases in BMD affect bone strength. For example, the implications of bone mass laid down on periosteum vs. endosteum may differ in relation to altering bone strength.

A second factor that warrants caution in the interpretation of our results is the relatively low quality of the studies included in our analysis. For example, only one study received a score of 3 while the remaining studies received scores ranging from 0 to 2. In addition, the different types of exercise interventions varied considerably. Furthermore, the fact that our funnel plot analysis was indicative of publication bias suggests that the results of this study may be an overestimate of the effects of exercise on BMD in men. Ideally, to examine the efficacy and effectiveness of exercise on BMD in men, it is suggested that future studies *1*) randomize subjects to an exercise and control condition, *2*) blind the person responsible for the assessment of BMD to the treatment assignment of subjects, *3*) limit participation to only subjects who have been previously sedentary, *4*) include weight-bearing exercise protocols that are reflective of those in which the population will be able to participate, *5*) assess BMD at the sites that were loaded during the exercise protocol, and *6*) include efficacy as well as effectiveness (intention-to-treat) analysis. In addition, it would be beneficial to examine any changes in BMD that might occur on cessation of exercise. Furthermore, it is critical that authors submit and be allowed to publish well-designed studies that yield null results. Consequently, we will be able to form a more valid conclusion regarding the true effects of exercise on BMD in men. Given the prevalence of low BMD in men $\geq$50 yr of age, it may be especially important to focus on this population (28).

Another factor that warrants caution in the interpretation of our results was the availability of data. Although meta-analysis is more quantitative than traditional narrative reviews, potential problems exist. The meta-analytic review, like any review, is limited by the available data. One potential problem in this investigation was the small number of ES available for some of the subgroup analyses. For example, ES changes in BMD at the specific sites we were able to examine (femur, lumbar, os calcis) were limited to two or three outcomes per site.

Despite our resampling approach because of the small sample sizes available for many of the analyses, additional studies in this area are needed before any firm conclusions can be made regarding the efficacy and effectiveness of exercise as a nonpharmacological intervention for improving and/or maintaining BMD in men. In addition, insufficient information was available to examine results according to the impact of the exercise protocols on the BMD sites assessed. Future studies need to provide a complete description of the exercise intervention.

In conclusion, the results of this study suggest that site-specific exercise may help increase and/or maintain BMD at the femur, lumbar, and os calcis sites in older men. However, the biological importance of the relatively small changes observed for most outcomes, quality of studies included, and limited data pool prevent us from forming any firm conclusion regarding the use of exercise for maintaining and/or improving BMD in men.

This study was funded by US Department of Defense, Army Medical Research and Material Command Award 17-98-1-8513.

Address for reprint requests and other correspondence: G. A. Kelley, Meta-Analytic Research Group, Dept. of Kinesiology and Physical Education, Anderson Hall, Rm. 233, Northern Illinois University, DeKalb, IL 60115-2854 (E-mail: gkelley@niu.edu).

Received 27 July 1999; accepted in final form 10 December 1999.

REFERENCES

1. **Begg CB.** Publication bias. In: *The Handbook of Research Synthesis*, edited by Cooper H and Hedges LV. New York: Russell Sage, 1994, p. 399–409.
2. **Bennell KL, Malcolm SA, Khan KM, Thomas SA, Reid RJ, Brukner PD, Ebeling PR, and Wark JD.** Bone mass and bone turnover in power athletes and controls: a 12-month longitudinal study. *Bone* 20: 477–484, 1997.
3. **Berard A, Bravo G, and Gauthier P.** Meta-analysis of the effectiveness of physical activity for the prevention of bone loss in postmenopausal women. *Osteoporos Int* 7: 331–337, 1997.
4. **Blumenthal JA, Emery CF, Madden DJ, Schniebolk S, Riddle MW, Cobb FR, Higgenbotham M, and Coleman RE.** Effects of exercise training on bone density in older men and women. *J Am Geriatr Soc* 39: 1065–1070, 1991.
5. **Braith RW, Mills RM, Welsch MA, Keller JW, and Pollock ML.** Resistance training restores bone mineral density in heart transplant patients. *J Am Coll Cardiol* 28: 1471–1477, 1996.
6. **Cohen B, Millet PJ, Mist B, Laskey MA, and Rushton N.** Effect of exercise training programme on bone mineral density in novice college rowers. *Br J Sports Med* 29: 85–88, 1995.
7. **Cohen J.** *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic, 1988, p. 1–17.
8. **Cooper HC and Hedges LV** (Editors). *The Handbook of Research Synthesis*. New York: Russell Sage, 1994, p. 3–14.
9. **Dickersin K, Hewett P, Mutch L, Chalmers I, and Chalmers TC.** Perusing the literature: comparison of Medline searching with a perinatal trials database. *Control Clin Trials* 6: 306–317, 1985.
10. **Efron B and Tibshirani R.** *An Introduction to the Bootstrap*. London: Chapman & Hall, 1993, p. 10–15.
11. **Egger M and Davey Smith G.** Meta-analysis: bias in location and selection of studies. *Br Med J* 316: 61–66, 1998.
12. **Follmann D, Elliot P, Suh I, and Cutler J.** Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 45: 769–773, 1992.
13. **Fujimura R, Ashizawa N, Watanabe M, Mukai N, Amagai H, Fukubayashi T, Hayashi T, Tokuyama K, and Suzuki M.** Effect of resistance exercise training on bone formation and resorption in young male subjects assessed by biomarkers of bone metabolism. *J Bone Miner Res* 12: 656–662, 1997.
14. **Glass GV, McGaw B, and Smith ML.** *Meta-Analysis in Social Research*. Newbury Park, CA: Sage, 1981, p. 21–24.
15. **Hedges LV and Olkin I.** Vote-counting methods in research synthesis. *Psychol Bull* 88: 359–369, 1980.
16. **Hedges LV and Olkin I.** *Statistical Methods for Meta-Analysis*. San Diego, CA: Academic, 1985, p. 13, 78, 122–130, 149–165, 167–183.
17. **Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, and McQuay HJ.** Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 17: 1–12, 1996.
18. **Jones G, Nguyen T, Sambrook PN, Kelly PJ, Gilbert C, Eisman JA.** Symptomatic fracture incidence in elderly men and women: the Dubbo Osteoporosis Epidemiology Study (DOES). *Osteoporos Int* 4: 277–282, 1994.
19. **Kelley GA.** Aerobic exercise and bone density at the hip in postmenopausal women: a meta-analysis. *Prev Med* 27: 798–807, 1998.
20. **Kelley GA.** Aerobic exercise and lumbar spine bone mineral density in postmenopausal women: a meta-analysis. *J Am Geriatr Soc* 46: 143–152, 1998.
21. **Kelley GA.** Exercise and regional bone mineral density in postmenopausal women: a meta-analytic review of randomized trials. *Am J Phys Med Rehabil* 77: 76–87, 1998.
22. **Light RJ and Pillemer DB.** *Summing Up: The Science of Reviewing Research*. Cambridge, MA: Harvard University Press, 1984, p. 65–69.
23. **McCartney N, Hicks AL, Martin J, and Webber CE.** Long-term resistance training in the elderly: effects on dynamic strength, exercise capacity, muscle, and bone. *J Gerontol A Biol Sci Med Sci* 50: B97–B104, 1995.
24. **Menkes A, Mazel S, Redmond RA, Koffler K, Libanati CR, Gundberg CM, Zizic TM, Hagberg JM, Pratley RE, and Hurley BF.** Strength training increases regional bone mineral density and bone remodeling in middle-aged and older men. *J Appl Physiol* 74: 2478–2484, 1993.
25. **Michel BA, Lane NE, Bjorkengren A, Bloch DA, and Fries JF.** Impact of running on lumbar bone density: a 5-year longitudinal study. *J Rheumatol* 19: 1759–1763, 1992.
26. **Mosteller F and Colditz GA.** Understanding research synthesis (meta-analysis). *Annu Rev Public Health* 17: 1–23, 1996.
27. **National Center for Health Statistics.** *Healthy People 2000 Review, 1997*. Hyattsville, MD: US Public Health Service, 1997.
28. **National Osteoporosis Foundation.** *1996 and 2015 Osteoporosis Prevalence Figures: State by State Report*. Washington, DC: Natl. Osteoporos. Found., 1997.

28a. **Nguyen TV, Eisman JA, Kelley PJ, and Sambrook PN** Risk factors for osteoporotic fractures in elderly men. *Am J Epidemiol* 144: 255–263, 1996.

29. **Petitti DB.** *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. New York: Oxford University Press, 1994, p. 15–20.
30. **Potvin C and Roff D.** Distribution-free and robust statistical methods: viable alternatives to parametric statistics? *Ecology* 74: 1617–1628, 1993.
31. **Pritchard JE, Nowson CA, and Wark JD.** Bone loss accompanying diet-induced or exercise-induced weight loss: a randomised controlled study. *Int J Obes* 20: 513–520, 1996.
32. **Recker RR, Davies M, Hinders SM, Heaney RP, Stegman MR, and Kimmel DB.** Bone gain in young adult women. *JAMA* 268: 2403–2408, 1992.
33. **Riggs BL and Melton LJ.** The worldwide problem of osteoporosis: insights afforded by epidemiology. *Bone* 17, *Suppl*: 505S–511S, 1995.
34. **Ryan AS, Treuth MS, Rubin MA, Miller JP, Nicklas BJ, Landis DM, Pratley RE, Libanati CR, Gundberg CM, and Hurley BF.** Effects of strength training on bone mineral density: hormonal and bone turnover relationships. *J Appl Physiol* 77: 1678–1684, 1994.
35. **Snow-Harter C and Marcus R.** Exercise, bone mineral density, and osteoporosis. *Exerc Sport Sci Rev* 10: 351–388, 1991.
36. **Williams JA, Wagner J, Wasnich R, and Heilbrun L.** The effect of long-distance running upon appendicular bone mineral content. *Med Sci Sports Exerc* 16: 223–227, 1984.
37. **Wolff I, van Croonenborg JJ, Kemper HCG, Kostense PJ, and Twisk JWR.** The effect of exercise training programs on bone mass: a meta-analysis of published controlled trials in pre- and postmenopause women. *Osteoporos Int* 9: 1–12, 1999.
38. **Zhu W.** Making bootstrap statistical inferences: a tutorial. *Res Q Exerc Sport* 68: 44–55, 1997.