frontiers in Physiology

SYSTEMATIC REVIEW
published: 02 December 2021
doi: 10.3389/fphys.2021.782695

Systematic Review and Meta-Analysis of Endurance Exercise Training Protocols for Mice

Michael P. Massett, Caitlyn Matejka and Hyoseon Kim*

Department of Kinesiology and Sport Management, Texas Tech University, Lubbock, TX, United States

Inbred and genetically modified mice are frequently used to investigate the molecular mechanisms responsible for the beneficial adaptations to exercise training. However, published paradigms for exercise training in mice are variable, making comparisons across studies for training efficacy difficult. The purpose of this systematic review and meta-analysis was to characterize the diversity across published treadmill-based endurance exercise training protocols for mice and to identify training protocol parameters that moderate the adaptations to endurance exercise training in mice. Published studies were retrieved from PubMed and EMBASE and reviewed for the following inclusion criteria: inbred mice; inclusion of a sedentary group; and exercise training using a motorized treadmill. Fifty-eight articles met those inclusion criteria and also included a "classical" marker of training efficacy. Outcome measures included changes in exercise performance, $\dot{V}O_{2max}$, skeletal muscle oxidative enzyme activity, blood lactate levels, or exercise-induced cardiac hypertrophy. The majority of studies were conducted using male mice. Approximately 48% of studies included all information regarding exercise training protocol parameters. Meta-analysis was performed using 105 distinct training groups (i.e., EX-SED pairs). Exercise training had a significant effect on training outcomes, but with high heterogeneity (Hedges' $g = 1.70$, 95% CI = 1.47–1.94, $Tau^2 = 1.14$, $I2 = 80.4\%$, prediction interval = –0.43–3.84). Heterogeneity was partially explained by subgroup differences in treadmill incline, training duration, exercise performance test type, and outcome variable. Subsequent analyses were performed on subsets of studies based on training outcome, exercise performance, or biochemical markers. Exercise training significantly improved performance outcomes (Hedges' $g = 1.85$, 95% CI = 1.55–2.15). Subgroup differences were observed for treadmill incline, training duration, and exercise performance test protocol on improvements in performance. Biochemical markers also changed significantly with training (Hedges' $g = 1.62$, 95% CI = 1.14–2.11). Subgroup differences were observed for strain, sex, exercise session time, and training duration. These results demonstrate there is a high degree of heterogeneity across exercise training studies in mice. Training duration had the most significant impact on training outcome. However, the magnitude of the effect of exercise training varies based on the marker used to assess training efficacy.

Keywords: inbred mice, treadmill running, sedentary, training responses, endurance exercise training

OPEN ACCESS

Edited by:
Hamdi Chtourou,
University of Sfax, Tunisia

Reviewed by:
Junichi Suzuki,
Hokkaido University of Education, Japan
Xu Yan,
Victoria University, Australia, Australia

****Correspondence:***
Michael P. Massett
mmassett@ttu.edu

Specialty section:
This article was submitted to Exercise Physiology, a section of the journal Frontiers in Physiology

Received: *24 September 2021*
Accepted: *05 November 2021*
Published: *02 December 2021*

Citation:
Massett MP, Matejka C and Kim H (2021) Systematic Review and Meta-Analysis of Endurance Exercise Training Protocols for Mice. Front. Physiol. 12:782695. doi: 10.3389/fphys.2021.782695

INTRODUCTION

Cardiovascular disease accounts for one in four deaths (~23%) in the United States (Murphy et al., 2021). Cancer, Alzheimer's disease, diabetes, and hypertension also rank within the top 15 causes of death in the United States (Murphy et al., 2021). These chronic diseases have been linked to low levels of cardiorespiratory fitness (Defina et al., 2013; Zhang et al., 2014; Zaccardi et al., 2015; Sui et al., 2017; Robsahm et al., 2019; Lee, 2021). The Centers for Disease Control and Prevention and the American College of Sports Medicine currently recommend that individuals participate in moderate-intensity physical activity for 150 or more minutes per week for optimal health (Haskell et al., 2007; Garber et al., 2011). Improving cardiorespiratory fitness through increased physical activity can significantly reduce the risk of all-cause mortality (Blair et al., 1995; Brawner et al., 2017; Davidson et al., 2018). Although the majority of health benefits associated with high cardiorespiratory fitness are mediated by changes in traditional risk factors such as blood pressure, inflammatory markers, and blood lipids, roughly 40% of the beneficial effects of exercise cannot be explained by traditional risk factors (Mora et al., 2007; Joyner and Green, 2009). Furthermore, the cellular and molecular mechanisms underlying the salutary effects of exercise are not well understood. Therefore, inbred and genetically modified mice are frequently used to investigate the integrative physiological responses to exercise and the molecular mechanisms responsible for the beneficial adaptations to exercise training.

There are three commonly used paradigms for exercise training in rodents – swimming, voluntary wheel running, and "forced" wheel or treadmill running – and each has been used to study the molecular basis of responses to acute exercise and chronic exercise training. Treadmill running and wheel running induce adaptations in mice associated with endurance exercise training (Allen et al., 2001; Kemi et al., 2002; De Angelis et al., 2004; Waters et al., 2004; Massett and Berk, 2005; Chow et al., 2007). However, the two paradigms are inherently different (Poole et al., 2020) such that the correlation between treadmill running performance and voluntary wheel-running performance among mouse strains is nominal (Allen et al., 2001; Lightfoot et al., 2001, 2004). One advantage of treadmill running as an exercise paradigm is that the total amount of work performed among all mice can be established by the investigator through the selection of exercise testing and training parameters. Unlike for humans, there are no published well-accepted standards for exercise training paradigms or levels of activity required for optimal changes in exercise capacity or other training adaptations (Fuller and Thyfault, 2021). The published exercise testing and training paradigms are quite variable (Kemi et al., 2002; Billat et al., 2005; Hoydal et al., 2007; Marcaletti et al., 2011; Ayachi et al., 2016; Petrosino et al., 2016). Therefore, the purpose of this review was to characterize the variation in exercise training protocols in mice and determine key training parameters involved in adaptations to exercise training. This review focuses on treadmill running because the training parameters can be more easily quantified and any potential recommendations regarding these parameters could be incorporated into future research utilizing treadmill-based exercise training in mice.

MATERIALS AND METHODS

The protocol for systematic reviews of animal studies was used as a guide for this review and meta-analysis (de Vries et al., 2015).

The following terms were used to search PubMed and EMBASE databases: (((((((""Inbred Mouse Strains"" OR ""Inbred Strain of Mice"" OR ""Inbred Strain of Mouse"" OR ""Inbred Strains of Mice"" OR ""Mice, Inbred Strains""[MeSH Terms] OR ""Mice, Inbred Strains"" OR ""Mouse, Inbred Strain"") OR (""Mice""[MeSH Terms] OR ""Mice"" OR ""Mice, House"" OR ""Mice, Laboratory"" OR ""Mouse"" OR ""Mouse, House"" OR ""Mouse, Laboratory"" OR ""Mouse, Swiss"" OR ""Mus"" OR ""Mus domesticus"" OR ""Mus musculus"" OR ""Mus musculus domesticus"" OR ""Swiss Mice""))) NOT ((transgenic OR knockout OR db/db OR ob/ob OR mdx OR ApoE)))) AND (((((""Aerobic Exercise"" OR ""Exercise""[MeSH Terms] OR ""Exercise"" OR ""Exercise Training"" OR ""Exercise, Aerobic"" OR ""Exercise, Physical"" OR ""Physical Activity"" OR ""Running""[MeSH Terms] OR ""Running""))) OR (((""High-Intensity Intermittent Exercise"" OR ""High-Intensity Interval Training""[MeSH Terms] OR ""High-Intensity Interval Training"" OR ""Sprint Interval Training""))))))) AND treadmill) AND (sedentary OR control)." Additional abstracts were obtained from reference lists of potentially eligible articles. The search was competed in February 2020.

Inclusion and Exclusion Criteria

Studies were included if they utilized inbred or wild-type mice of any strain divided into at least two groups: exercise training and sedentary control, the duration of the exercise training protocol was at least 1 week, and the training was performed on a motor-driven treadmill. Studies also needed to include an outcome measure of training efficacy reported for both the exercise-trained and sedentary control groups. Acceptable outcome measures included assessment of exercise performance or oxygen consumption ($\dot{V}O_2$), skeletal muscle oxidative enzymes (e.g., citrate synthase), post-exercise blood lactate levels, skeletal muscle fiber types, or other markers of metabolic or cardiovascular adaptation (Holloszy and Coyle, 1984; Booth et al., 2010; Hellsten and Nyberg, 2015). If studies reported more than one outcome variable, performance outcomes based on the results of an exercise performance test were prioritized over other outcomes (Vesterinen et al., 2014). Studies that involved mice receiving a treatment other than exercise on a treadmill such as a diet or drug intervention were excluded. For studies that included four or more groups of mice – a control arm combined with exercise training (e.g., no treatment ± exercise training) and a treatment arm combined with exercise training (treatment ± exercise training), only the mice in the control arm were included in the analysis. Genetic manipulation or modification can have a significant impact on exercise performance. Therefore, this review focused on inbred or wild-type mice of any strain.

Studies utilizing only transgenic or genetically manipulated mice were excluded as were mice performing swimming, wheel running, or other forms of exercise training. Several studies utilized a treadmill-based overtraining paradigm. Because this paradigm generally resulted in decreased performance, cohorts undergoing overtraining were excluded. However, if a traditional exercise training paradigm was included as part of the study and efficacy data available, data from mice in those cohorts were included. Any studies involving other animals or humans were excluded as were studies that did not report sufficient training efficacy data.

Study Selection and Data Extraction

Following the initial search, titles and abstracts were screened for (1) inbred mice with no treatment; (2) inclusion of a sedentary/control group; (3) exercise training; and (4) training with a treadmill. Full-text articles were then assessed against the inclusion criteria. Data extracted included: author names, publication date and journal citation, sex and age of the mice, number of mice per group, exercise training protocol variables – frequency (days/week), session duration (min), treadmill velocity (m/min), treadmill incline (degrees), training duration (weeks), intensity (% of maximum), type of exercise performance test, and exercise training efficacy outcome variables for each group. In studies where the exercise protocol progressively increased to a maximal target workload, the final workload was used in all analyses. In some studies, the subject characteristics (e.g., age) or final training protocol variables (e.g., treadmill velocity) were presented as a range. In those cases, the median value was used for all analyses. The mean and standard deviation (SD) or standard error of the mean (SEM) were recorded for each outcome variable. If the outcome data were presented in figures, data were extracted using WebPlotDigitizer.[1] Two investigators extracted data independently. A third investigator reviewed the data, calculated the average, or requested a re-analysis by both investigators.

Quality Assessment of Included Studies

Risk of bias was assessed using a modified version of the CAMARADES checklist items (Macleod et al., 2004). The following reported items were recorded: (1) random assignment to groups, (2) blinded assessment of outcome variables, (3) sample size calculation, (4) animal welfare statement, and (5) conflict of interest statement.

[1]https://automeris.io/WebPlotDigitizer/

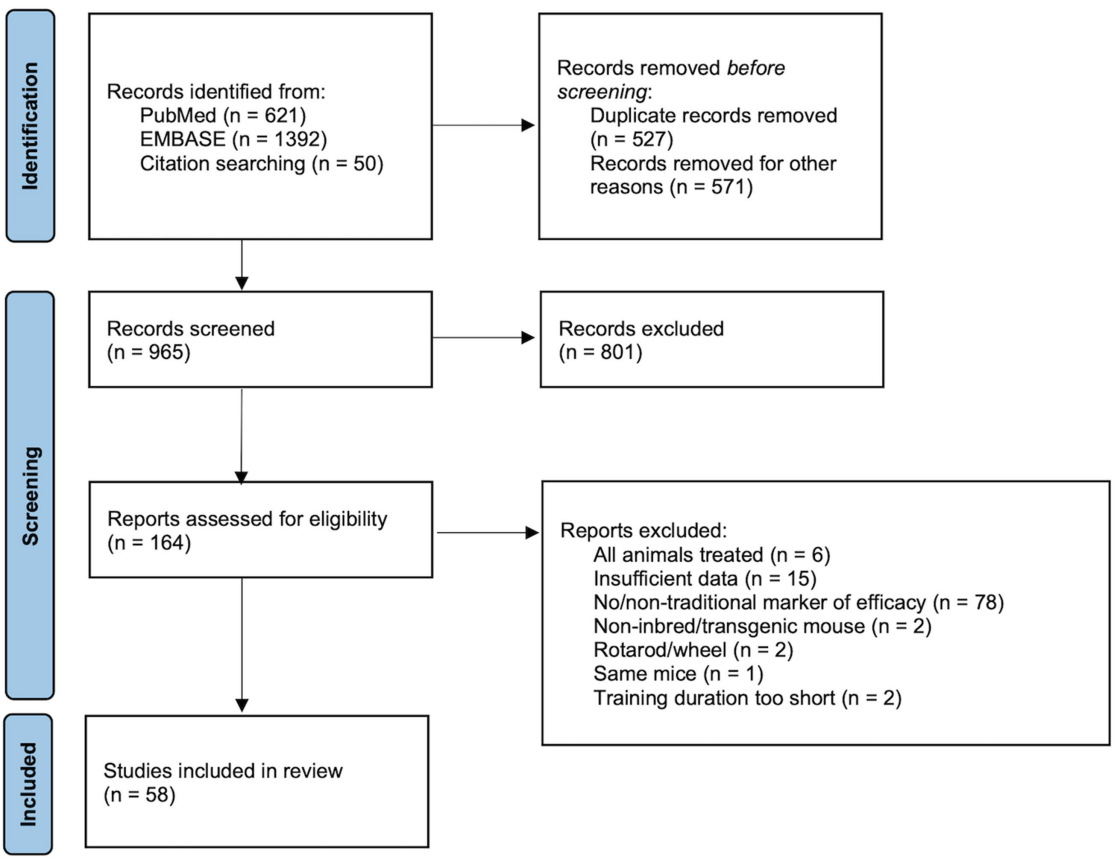


FIGURE 1 | Flow chart of article selection process.

TABLE 1 | Summary of mouse characteristics and training parameters from studies included in meta-analysis.

Study	Subject Characteristics	Training Protocol	Outcome
Abadi et al., 2013	*Strain(s):* C57BL/6J *Sex:* Female and Male *Age:*	*Frequency:* *Velocity:* 16–18 m/min *Incline:* *Session duration:* 45 min *Training duration:* 8 weeks *Intensity:*	Distance (m), Incremental load test
Aguiar et al., 2008	*Strain(s): CF*-1 *Sex:* Male *Age:* 6 weeks	*Frequency:* 5 d/wk. *Velocity:* 16.5 m/min *Incline:* 0° (Level group) *Session duration:* 45 min *Training duration:* 8 weeks *Intensity:*	Citrate synthase activity ($mmol \bullet min^{-1} \bullet g^{-1}$), soleus
Almeida-Oliveira et al., 2019	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* 50% maximal exercise capacity	Time (min), Incremental load test
Alves et al., 2017	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 13 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 7 weeks *Intensity:* 55–65% of maximal speed	Distance (m), Incremental load test
Alves et al., 2019	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 13 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 12 weeks *Intensity:* 70% maximal exercise capacity	Work (Joules), Incremental load test
Alves et al., 2020	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 14° *Session duration:* 60 min *Training duration:* 12 weeks *Intensity:* 60% of maximal speed	% change in velocity, Incremental load test
Avila et al., 2017	*Strain(s):* 24 strains *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* 15–23 m/min *Incline:* 5–10° *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* 65%	Change in time (min), graded exercise test
Bartalucci et al., 2012	*Strain(s):* C57BL *Sex:* Male *Age:* 10 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* *Training duration:* *Intensity:* HIT: 2 min @ 90% max, 1 min recovery to 1,000 meters, LOW: 60% of maximal velocity to 1,000 meters	Blood lactate concentration ($mmol \cdot L^{-1}$), at the end of the first and last training session

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
Boehnke et al., 1987	*Strain(s):* Swiss Webster *Sex:* Female *Age:* 9.5 weeks	*Frequency:* 6 d/wk. *Velocity:* moderate: 7 m/min, high: 15 m/min *Incline:* 6° *Session duration:* 60 min *Training duration:* 9 weeks *Intensity:*	Succinate dehydrogenase activity (umol/g tissue x min), gastrocnemius
Borg et al., 2014	*Strain(s):* C57BL/6J *Sex:* Male *Age:* 12 weeks	*Frequency:* 5 d/wk. *Velocity:* 18 m/min *Incline:* 5° *Session duration:* 70 min *Training duration:* 6 weeks *Intensity:*	Time (s), Incremental load test
De Angelis et al., 2004	*Strain(s):* C57/6J *Sex:* Male *Age:*	*Frequency:*5 d/wk. *Velocity:* 17 m/min *Incline:* *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* 50–70% maximal running speed	Speed (km/h), graded exercise test
Durigan et al., 2009a	*Strain(s):* BALB/c *Sex:* Male *Age:* 16 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 12 weeks *Intensity:* 50% maximal speed and 75% maximal speed	Improvement in maximal exercise capacity (min), incremental load test
Durigan et al., 2009b	*Strain(s):* BALB/c *Sex:* Male *Age:* 16 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 12 weeks *Intensity:* 50% maximal speed and 75% maximal speed	Time (min), Incremental load test
Ferreira et al., 2007	*Strain(s):* C57BL/6J *Sex:* Male *Age:* 20 weeks	*Frequency:* 5 d/wk. *Velocity:* 15.1 m/min *Incline:* *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:* MLSSw	Total distance run (m), Incremental load test
Fiuza-Luces et al., 2018	*Strain(s):* C57BL/6J *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 8.5° *Session duration:* 50 min *Training duration:* 8 weeks *Intensity:* 70–75% Vmax	Total running distance (m), Incremental load test
Foryst-Ludwig et al., 2011	*Strain(s):* C57BL/6J *Sex:* Female, Male *Age:* 5 weeks	*Frequency:* 7 d/wk. *Velocity:* 15 m/min *Incline:* 7° *Session duration:* 90 min *Training duration:* 3 weeks progressive increase and 4 weeks at final workload *Intensity:*	Blood lactate (mg/dl)

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
German and Hoffman-Goetz, 1986	*Strain(s):* C57BL/6J *Sex:* Male *Age:* 72–80 weeks	*Frequency:* 5 d/wk. *Velocity:* 15 m/min *Incline:* 0° *Session duration:* 30 m/min *Training duration:* 8 weeks *Intensity:*	Succinate dehydrogenase activity (μmoles/gm protein/min), vastus intermedius
Han, 2013	*Strain(s):* 129 SvJ/C57BL6 *Sex:* Male *Age:* 15 weeks	*Frequency:* 5 d/wk. *Velocity:* 24 m/min *Incline:* *Session duration:* 40 min *Training duration:* 8 weeks *Intensity:*	Ratio of heart weight (mg) to body weight (g)
Herbst et al., 2015	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 7 weeks	*Frequency:* 5 d/wk. *Velocity:* 25 m/min *Incline:* 11.3° *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:*	mtDNA copy number (vastus lateralis)
Hoffman-Goetz et al., 1986	*Strain(s):* C57BL/6J *Sex:* Male *Age:* 12 weeks	*Frequency:* 5 d/wk. *Velocity:* 28 m/min *Incline:* 8° *Session duration:* 30 min *Training duration:* 4 weeks *Intensity:*	Succinate dehydrogenase activity (μmoles/g protein/min), quadriceps femoris
Hoffman-Goetz et al., 1989	*Strain(s):* C3He *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* 30 m/min *Incline:* 8° *Session duration:* 30 min *Training duration:* 8 weeks *Intensity:*	Succinate dehydrogenase activity (μmoles/g protein/min), quadriceps femoris
Ingalls et al., 1996	*Strain(s):* ICR *Sex:* Male *Age:* 7 weeks	*Frequency:* 4 d/wk. *Velocity:* 27–36 m/min *Incline:* 9.9° *Session duration:* 3 sets of 3 min (30 s recovery) *Training duration:* 8 weeks *Intensity:*	Work output (joules), Incremental load test
Jadeski and Hoffman-Goetz, 1996	*Strain(s):* C3H/HeJ *Sex:* *Age:* 4–9 weeks	*Frequency:* 5 d/wk. *Velocity:* 20 m/min *Incline:* 0° *Session duration:* 30 min *Training duration:* 9 weeks *Intensity:*	Citrate synthase activity (mol/min/g tissue), soleus
Kaurstad et al., 2012	*Strain(s):* C57BL/6J *Sex:* Female *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 25° *Session duration:* 60 min *Training duration:* 6 weeks *Intensity:* 4 min @ 85–90% $\dot{V}O_{2max}$, 2 min @ 50–60% $\dot{V}O_{2max}$	Maximal oxygen uptake ($\dot{V}O_{2max}$) (ml/kg$^{0.75}$/min)

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
Kemi et al., 2002	*Strain(s):* C57BL/6J *Sex:* Female and Male *Age:* 7–8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 25° *Session duration:* 120 min *Training duration:* 8 weeks *Intensity:* 8 min @ 85–90% $\dot{V}O_{2max}$, 2 min @ 50–60% $\dot{V}O_{2max}$	Maximal oxygen uptake ($\dot{V}O_{2max}$) (ml/kg$^{0.75}$/min)
Kim et al., 2019	*Strain(s):* ICR *Sex:* Male *Age:* 7 weeks	*Frequency:* 5 d/wk. *Velocity:* 18 m/min *Incline:* 8° *Session duration:* 50 min *Training duration:* 4 weeks *Intensity:* 60% $\dot{V}O_{2max}$	Oxygen uptake during 1 h of exercise (ml/kg/min)
Kim et al., 2020	*Strain(s):* C57BL/6J, 129S1/SvlmJ, SJL/J, NON/ShiLtJ *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* HIT: 60 min, MOD: 70 min *Training duration:* 4 weeks *Intensity:* HIT: 8 min @ 85% max speed, 2 min @ 50% max speed; MOD: 65% maximal speed	Change in time (min), Graded exercise test
Kruger et al., 2013	*Strain(s):* C57BL/6N *Sex:* Male *Age:* 10–12 weeks	*Frequency:* 5 d/wk. *Velocity:* 12 m/min *Incline:* *Session duration:* 35 min *Training duration:* 10 weeks *Intensity:*	Maximal oxygen consumption ($\dot{V}O_{2max}$) (ml/min/kg)
Kruger et al., 2016	*Strain(s):* C57BL/6N *Sex:* Male *Age:* 10–12 weeks	*Frequency:* 5 d/wk. *Velocity:* 15.6 m/min *Incline:* *Session duration:* 35 min *Training duration:* 10 weeks *Intensity:* 80% $\dot{V}O_{2max}$	Maximal oxygen consumption ($\dot{V}O_{2max}$) (ml/min/kg)
Lee et al., 2015	*Strain(s):* C57BL/6N *Sex:* Male *Age:* 56 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 5° *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:* 60% of maximum work rate	Total distance run (m), Incremental load test
Lehti et al., 2006	*Strain(s):* NMRI *Sex:* Male *Age:* 10 weeks	*Frequency:* 5 d/wk. *Velocity:* 21 m/min *Incline:* 2.5° *Session duration:* 60 min *Training duration:* 5 weeks *Intensity:*	Citrate synthase activity (nmol•min^{-1} •mg^{-1}), calf muscle complex
Liu et al., 2008	*Strain(s):* BALB/c *Sex:* Male *Age:* 12 weeks	*Frequency:* 5 d/wk. *Velocity:* 10 m/min *Incline:* *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* moderate	Citrate synthase activity (μmol/min/mg protein), soleus

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
Lucchetti et al., 2017	*Strain(s):* Swiss *Sex:* Male *Age:*	*Frequency:* 5 d/wk. *Velocity:* 15.1 m/min *Incline:* *Session duration:* 60 min *Training duration:* 9 weeks *Intensity:*	Maximum speed (m/min), Incremental load test
Malek et al., 2013	*Strain(s):* FVB/NJ *Sex:* Male *Age:* 16 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 5° *Session duration:* 30 min continuous or 3 × 10 min (2 h recovery between) *Training duration:* 8 weeks *Intensity:* 60% of the maximal work rate	Time (seconds), Incremental load test
Massett and Berk, 2005	*Strain(s):* BALB/cJ, C57BL/6 J, FVB/NJ *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* B6, BALB: 15 m/min, FVB: 19 m/min *Incline:* B6, BALB: 5°, FVB: 10° *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* ~60% of the maximal workload	Time (min), Graded exercise test
Meier et al., 2013	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 9 weeks	*Frequency:* 5 d/wk. *Velocity:* 26 m/min *Incline:* 10° *Session duration:* 45 min *Training duration:* 4 weeks *Intensity:*	Time (seconds), Incremental load test
Mikami et al., 2004	*Strain(s):* ICR *Sex:* Male *Age:* 10 weeks	*Frequency:* 5 d/wk. *Velocity:* 25 m/min *Incline:* 0° *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:*	Citrate synthase activity (U/mg protein), soleus
Niebauer et al., 1999	*Strain(s):* C57BL/6 J *Sex:* Female *Age:* 8 weeks	*Frequency:* 6 d/wk. *Velocity:* 22 m/min *Incline:* 8° *Session duration:* 120 min (2 × 1 h/day) *Training duration:* 4 weeks *Intensity:* 85% of maximal oxygen uptake	Maximal oxygen consumption ($\dot{V}O_{2max}$) (ml/min/kg)
Niel et al., 2017	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 92 weeks	*Frequency: Group 1*: 5 d/wk., *Group 2*: 5 sessions over 2 weeks *Velocity: Group 1*: 14 m/min, *Group 2*: 3 m.min^{-2} × 11 min, 6 m.min^{-2} × 6 min, 12 m.min^{-2} × 3 min (30 min rest between) *Incline:* 0° *Session duration: Group 1*: 60 min continuous, *Group 2*: 20 min *Training duration: Group 1*: 4 weeks, *Group 2*: 2 weeks *Intensity:* 50% of the maximum running speed (Vpeak)	Time (min), Incremental load test

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
Pereira et al., 2012	*Strain(s):* Swiss *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 0° *Session duration:* 60 min *Training duration:* Trained: 8 weeks, Overtrained: 4 weeks (before overtraining protocol) *Intensity:* 60% of exhaustion velocity	Exhaustion time (min), Incremental load test
Pereira et al., 2013	*Strain(s):* Swiss *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 0° *Session duration:* 60 min *Training duration:* Trained: 8 weeks, Overtrained: 4 weeks (before overtraining protocol) *Intensity:* 60% of exhaustion velocity	Exhaustion time (min), Incremental load test
Pereira et al., 2014a	*Strain(s):* Swiss *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 0° *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:*	Percentage change between week 0 and week 8 for time to exhaustion, Incremental load test
Pereira et al., 2014b	*Strain(s):* Swiss *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 0° *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:*	Percentage change between week 0 and week 8 for exhaustion velocity, Incremental load test
Pereira et al., 2015	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 0° *Session duration:* 60 min *Training duration:* 4 weeks (before start of overtraining protocol) *Intensity:* 60% of exhaustion velocity	Exhaustion velocity (m/min), Incremental load test
Pereira et al., 2016	*Strain(s):* C57BL/6 *Sex:* Male *Age:*	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* 60% of maximal velocity	Maximal velocity (km/h), Incremental load test
Pinto et al., 2015	*Strain(s):* C57BL/6 N *Sex:* Male *Age:* 12 weeks	*Frequency:* 5 d/wk. *Velocity:* 15 m/min *Incline:* *Session duration:* 60 min *Training duration:* 6 weeks *Intensity:*	Systolic blood pressure (tail-cuff)
Rodrigues et al., 2019	*Strain(s):* C57BL/6 J *Sex:* Male *Age:* 20 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:* 60% of maximal speed	Total distance run (m), Incremental load test

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
Savage and McPherron, 2010	*Strain(s):* C57BL/6 *Sex:* Male *Age:* 12 weeks	*Frequency:* 5 d/wk. *Velocity:* 12 m/min *Incline:* 0° *Session duration:* 30 min *Training duration:* 4 weeks *Intensity:*	Citrate synthase activity (mu/μg protein), quadriceps
Sousa et al., 2019	*Strain(s):* C57BL6/JUnib *Sex:* Male *Age:* 6–7 weeks	*Frequency:* 5 d/wk. *Velocity:* *Incline:* 0° *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:* 10 min at 40% of maximal speed, 40 min at 50–60% of maximal speed, and 10 min at 40% of maximal speed	Time (min), Incremental load test
Steiner et al., 2011	*Strain(s):* ICR *Sex:* Male *Age:* 8 weeks	*Frequency:* 6 d/wk. *Velocity:* 25 m/min *Incline:* 2.9° *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:*	Time (min), Run to fatigue test
Sturgeon et al., 2015	*Strain(s):* C57BL/6 J *Sex:* Female *Age:* 8 weeks	*Frequency:* 5 d/wk. *Velocity:* 18 m/min *Incline:* *Session duration:* 60 min *Training duration:* 8 weeks *Intensity:*	Work (m•kg), Incremental load test
Suominen et al., 1980	*Strain(s):* NMRI *Sex:* Male *Age:* 3 weeks and 8 weeks	*Frequency:* 5 d/wk. *Velocity:* 18 m/min *Incline:* 5° *Session duration:* 80 min (2 × 40 min) *Training duration:* 4 weeks (3 wo mice), 3 weeks (8 wo mice) *Intensity:*	Dry weight of heart. (mg)
Toti et al., 2013	*Strain(s):* C57BL *Sex:* Male *Age:* 10 weeks	*Frequency:* 5 d/wk. *Velocity:* LOW: 17.1 m/min, HIT: 33.75 m/min *Incline:* *Session duration:* time to complete running 1,000 meters *Training duration:* 8 weeks *Intensity:* HIT: 90% of maximal running velocity for 2 min, 1 min recovery; LOW: 60% of maximal running velocity	Blood lactate concentrations ($mmol•L^{-1}$) before 1st training session and after last training session
Uddin et al., 2016	*Strain(s):* C57BL/6 J *Sex:* Female *Age:* 11 weeks at start of training	*Frequency:* 6 d/wk. *Velocity:* 15 m/min *Incline:* *Session duration:* 45 min *Training duration:* 6 weeks *Intensity:*	Citrate synthase activity (μmol/min/mg protein), quadriceps

(Continued)

TABLE 1 | Continued

Study	Subject Characteristics	Training Protocol	Outcome
Vieira et al., 2008	*Strain(s):* BALB/c *Sex:* Male *Age:*	*Frequency:* 5 d/wk. *Velocity:* *Incline:* *Session duration:* 60 min *Training duration:* 4 weeks *Intensity:* Low: 50% of maximal speed, Moderate: 75% of maximal speed	Time (min), Incremental load test
Vihko et al., 1979	*Strain(s):* NMRI *Sex:* Male *Age:* 9–11 weeks	*Frequency:* 5 d/wk. *Velocity:* 25 m/min *Incline:* 6° *Session duration:* 90 min *Training duration:* *Intensity:*	Citrate synthase activity (nmol substrate consumed/min per mg muscle), red Musculi quadriceps femoris (MQF)
Wernig et al., 1991	*Strain(s):* CBA/J *Sex:* Male *Age:* 12 weeks	*Frequency:* *Velocity:* 14 m/min *Incline:* 6° *Session duration:* 540 min (3 × 3 h with 30 min rest in the cages between the bouts) *Training duration:* 8–10 times at intervals of 3–5 days *Intensity:*	Total number of Type II fibers, soleus
Woods et al., 2003	*Strain(s):* BALB/c *Sex:* Male *Age:* 8 weeks and 72 weeks	*Frequency:* 5 d/wk. *Velocity:* 13–22 m/min *Incline:* *Session duration:* 45 min *Training duration:* 16 weeks *Intensity:*	Citrate synthase activity (μm g wet wt^{-1} min^{-1}), soleus

Data Analysis

All descriptive statistics were performed using JMP Pro 15 (SAS, Cary, NC, United States). Summary figures were generated using Prism 9 (GraphPad Software, La Jolla, CA, United States). All meta-analyses were conducted using Comprehensive Meta-Analysis Software v3 (Biostat Inc., Englewood, NJ). Statistical significance was set at $p < 0.05$. Outcome variables were reported as pre- and post-training values, post-training only values, or as changes in the outcome variable. Standardized mean difference values between exercise-trained and sedentary groups (exercise group minus control group) were calculated as Hedges' *g*. Positive values indicate an improvement with exercise training. If change score SD were not available, these were calculated using the study-specific correlation coefficient or a correlation coefficient of 0.6 between pre- and post-training values. The latter value is the mean of previously published correlation coefficients between pre- and post-training values for exercise performance phenotypes (Troxell et al., 2003; Massett and Berk, 2005; Avila et al., 2017). For studies with more than one exercise training group, the common control group was split into two groups with smaller samples sizes to avoid double counting of animals (Vesterinen et al., 2014). Standardized mean differences were calculated for each comparison and considered separate studies in all analyses. Thresholds were set as small, $|g| \leq 0.5$; medium, $|g| < 1.0$; large, $|g| < 1.5$; and very large, $|g| \geq 1.5$ (Labots et al., 2016).

To investigate the contribution of moderator variables on the effect of exercise training, study-level categorical and continuous variables were included individually and together in a random-effects meta-regression model. Categorical factors included strain, sex, exercise intensity, exercise performance test, and training outcome. Continuous variables included age, treadmill velocity and incline, frequency, time/session, and training duration.

Heterogeneity was evaluated using Cochran's *Q* test, I^2, and Tau^2. Prediction intervals were calculated using CMA prediction interval program.[2] Subgroup analysis was used to investigate the heterogeneity between the sample estimates based on study-level moderators: mouse strain, age, sex, outcome variable, exercise performance test type, and exercise training protocol variables.

To assess publication bias, the funnel plot of Hedges' *g* vs. standard error, Egger's regression, and Duval and Tweedie trim and fill were examined. Assuming a positive effect of exercise training on outcome variables, imputed missing studies were plotted to the left side of the mean.

[2]www.Meta-Analysis.com/Prediction

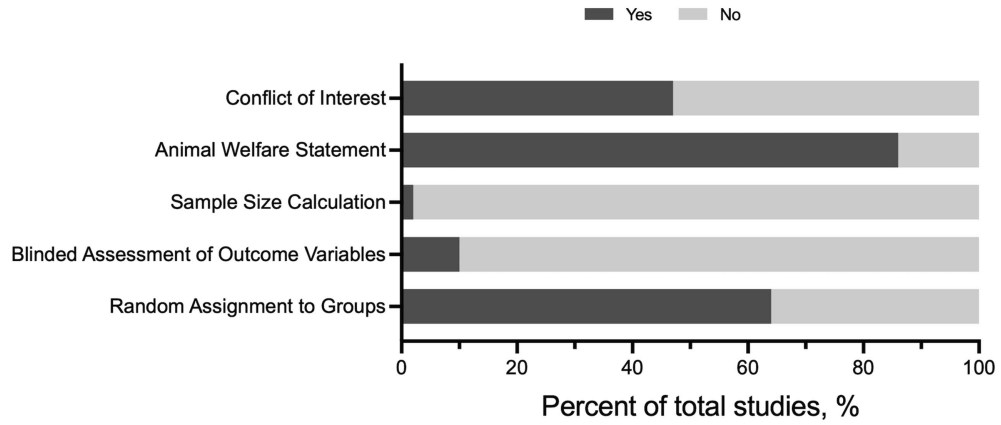


FIGURE 2 | Risk of bias assessment for 58 full-text articles meeting inclusion criteria.

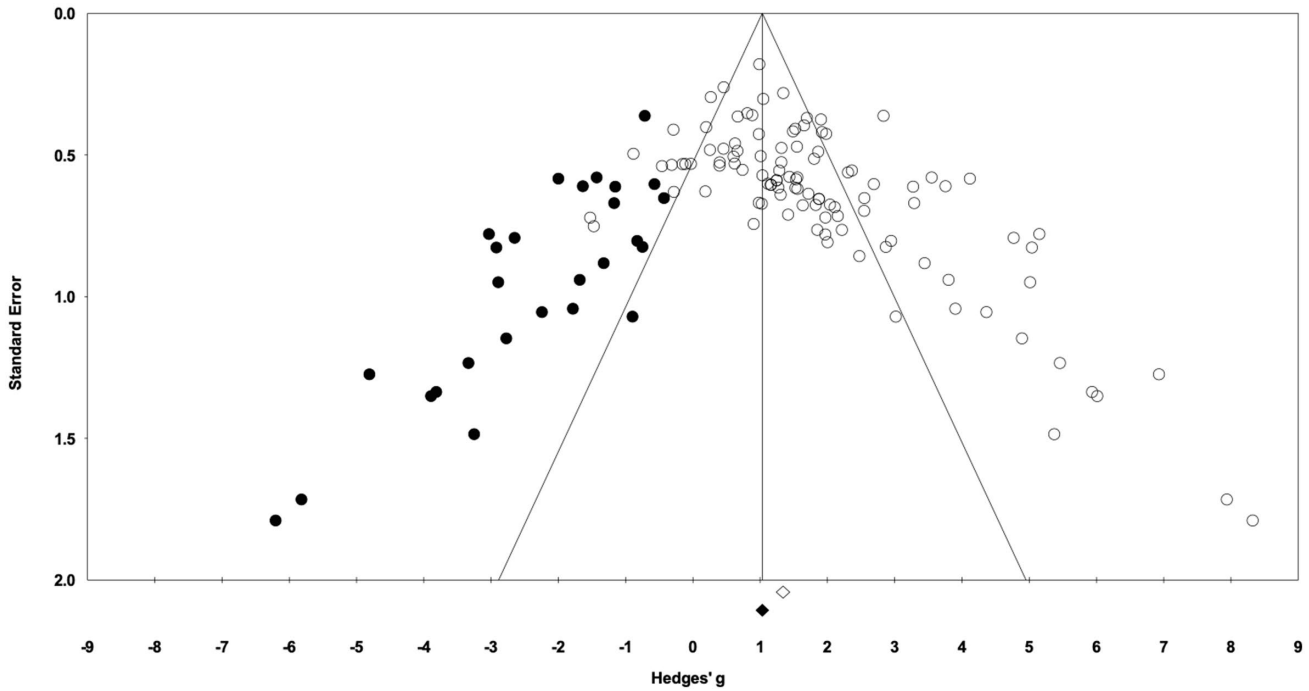


FIGURE 3 | Funnel plot of Hedges' *g* between exercise training and sedentary control groups. Open diamond indicates the point estimate and 95% CI for the combined studies using a random effects model. The black diamond indicates the point estimate based on the Duval and Tweedie's Trim and Fill analysis using a random effects model. Black circles are imputed studies from Trim and Fill analysis.

RESULTS

Selection Results

In the initial search, 2,063 articles were identified through database searches and other sources (i.e., reference lists and author publications). A flow chart based on PRISMA guidelines is shown in **Figure 1** Page et al. (2021)). Duplicate records ($n = 527$), non-full-text items ($n = 565$), and non-English language items ($n = 6$) were excluded. Of the remaining 965 articles, 801 articles were excluded based on the title and abstract review for: (1) inbred mice with no treatment, (2) inclusion of a sedentary/control group, (3) exercise training, and (4) training with a treadmill. The full text of 164 potentially eligible articles were assessed for inclusion criteria including markers of training responses. Of these, 106 articles were excluded for (1) utilizing transgenic/genetically modified mice, (2) including drug/diet supplement/treatment, (3) having different modalities of exercise such as running wheels and rotarod, or (4) no relevant exercise training phenotype. Fifty-eight (58) articles met the eligibility criteria and were included in the

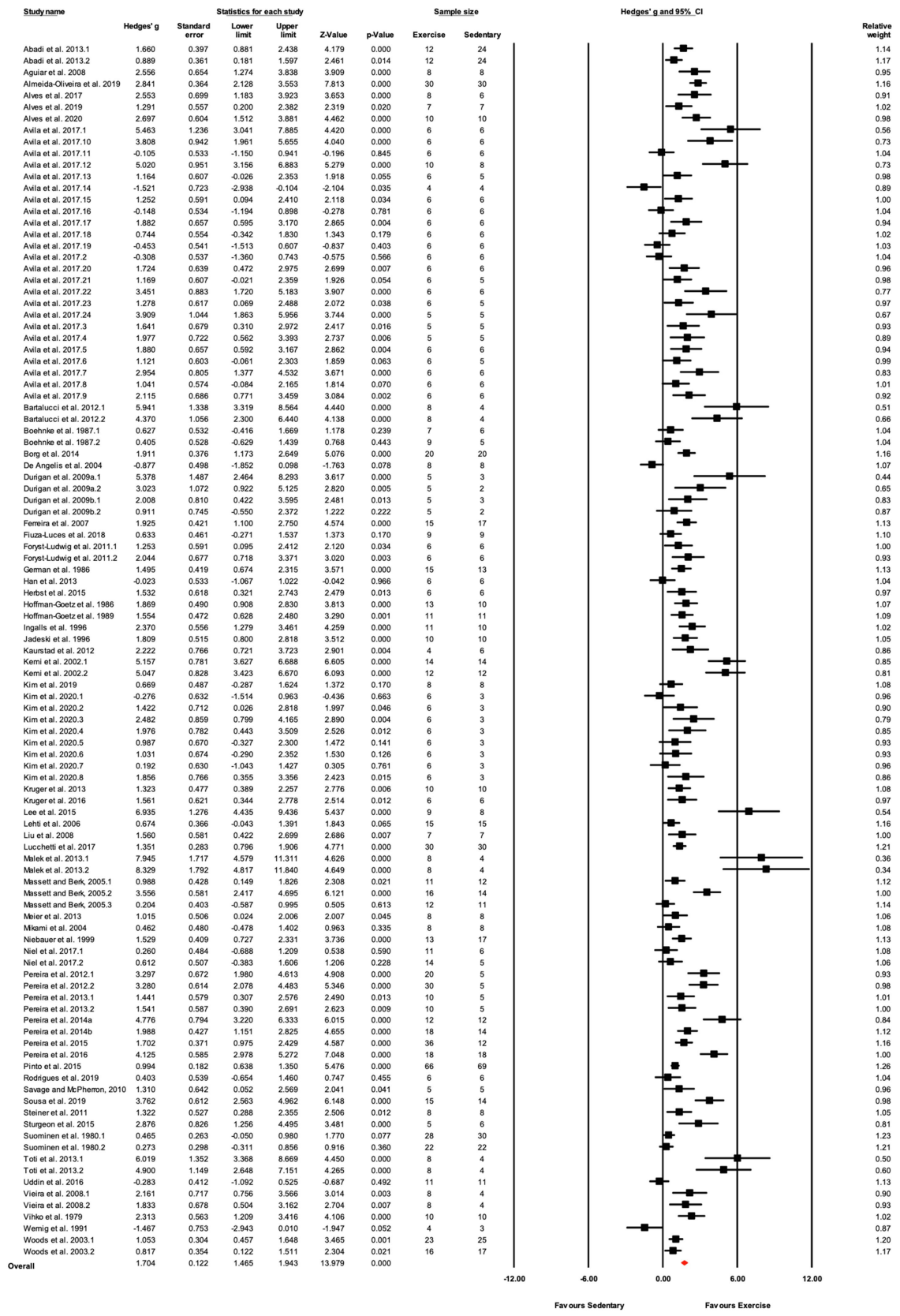


FIGURE 4 | Continued

FIGURE 4 | Mean difference effects of treadmill endurance training compared with sedentary control on markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Red diamond indicates pooled estimate of random effects model.

TABLE 2 | Subgroup analyses for the effect of exercise training on markers of training efficacy in mice.

Moderator variable	Subgroups	Number of studies	Between-group differences			Meta-regression
			Q-value	df	*p*	R^2 (%)
Strain	B6	41	2.77	1	ns	0
	Other	64				
Age	≤8 weeks	66	0.22	1	ns	0
	>8 weeks	32				
Sex	Male	95	0.0006	1	ns	0
	Female	9				
Training intensity	High	14	0.63	1	ns	0
	Moderate	58				
Treadmill velocity	≤10 m/min	3	3.97	2	ns	3
	11–20 m/min	43				
	>20 m/min	21				
Treadmill incline	≤5°	24	6.36	2	<0.05	2
	6–10°	37				
	>10°	5				
Time/session	≤30 min	11	2.40	3	ns	0
	31–45 min	10				
	46–60 min	66				
	>60 min	14				
Training duration	≤4 weeks	57	15.06	2	<0.05	0
	5–8 weeks	32				
	>8 weeks	14				
Performance test	GXT	36	11.34	2	<0.05	9
	ILT	36				
	$\dot{V}O_2$	6				
Outcome variable	Biochemical	20	17.58	2	<0.05	3
	Performance	78				
	Other	7				

GXT, graded exercise test; ILT, incremental load test; $\dot{V}O_{2max}$, maximal oxygen consumption test; df, degrees of freedom; ns, non-significant $p > 0.05$; R^2 (%), regression coefficient.

meta-analysis. Studies included in the systematic review and meta-analysis are summarized in **Table 1**.

Quality of the References and Publication Bias

Publication study quality was assessed in the full-text articles included for review (**Figure 2**). Out of 58 full-text articles, less than ~2% of the articles included a sample size calculation. Moreover, only six full-text articles (10%) included blinded assessment of the outcome. In general, the blinding assessment was applied to outcome variables not relevant to traditional exercise training-related phenotypes and not to group assignment or exercise performance tests. Random assignment to sedentary (SED) and exercise training (EX) groups was documented in 38 articles; however, the methods for randomization were not provided. The majority of the articles included an animal welfare statement (86%, 50/58 articles) and a conflict-of-interest statement (47%, 27/58). Collectively, these findings suggest there may be a potential bias in published articles, especially regarding appropriate sample size and blinding of outcomes.

In the 58 articles meeting the eligibility criteria, several reported results for more than one exercise training – sedentary cohort (e.g., multiple strains or sexes). For data analysis, each distinct training group (i.e., EX-SED pair) was considered a separate study, therefore, data from 105 studies are reported (i.e., 105 EX-SED pair comparisons). The assumption prior to starting this review was that there would be a significant bias toward the beneficial effects of exercise training; therefore, several approaches were used to assess publication bias. A significant effect of exercise training was observed in approximately 70% of included studies. The funnel plot in **Figure 3** shows the distribution of studies. A greater number of studies are located to the right of the mean effect size, suggesting some degree of publication bias. The random effects model point estimate and 95% CI for the combined studies was 1.70 (95% CI: 1.47–1.94). Using Trim and Fill the point estimate was 1.08 (95% CI: 0.82–1.35) with approximately 27 missing studies (**Figure 3**). The asymmetry was confirmed by the Egger's test. The intercept of the regression was 3.11 (95% CI: 2.11–4.12), with $t = 6.15$, df = 103, one-tailed value of $p < 0.05$.

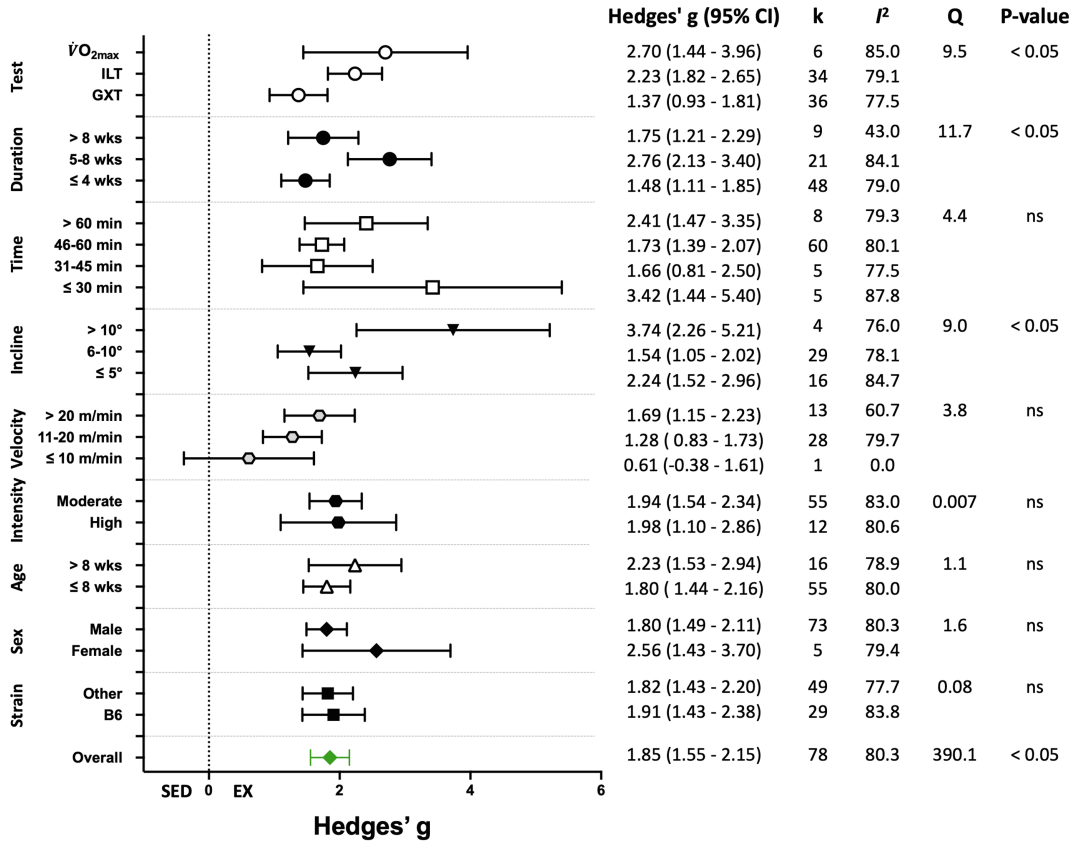


FIGURE 5 | Adjusted effect sizes of the between-group comparisons for the predefined moderators on performance-based exercise training outcomes. B6, C57BL/6 mice; GXT, graded exercise test; ILT, incremental load test; $\dot{V}O_{2max}$, maximal oxygen consumption test; k, number of studies in each subgroup; I^2, measure of heterogeneity; *Q*, Cochran's *Q*; *p*, value of *p* for heterogeneity analysis (overall) or differences between subgroups; ns, non-significant $p > 0.05$.

The result from the Egger's Test indicates significant asymmetry in the funnel plot (Egger et al., 1997).

Subject Characteristics

Data from 2,049 mice were reported in the 105 included studies. Twenty-eight different mouse strains were used in 58 full-text articles. C57BL/6 was the most used strain (39%), followed by BALB/c (10.5%), Swiss (9.5%), and NMRI, ICR, and FVB/NJ (3.8% each) strains. There was a marked difference in the number of studies that used male or female mice. Male mice were used exclusively in 88% of studies, whereas only a few studies (9%) utilized female mice. Three studies included both male and female mice (Kemi et al., 2002; Foryst-Ludwig et al., 2011; Abadi et al., 2013). The median age of mice was 8 weeks old with a range of 3–92 weeks old (mean ± SD, 13 ± 16 weeks) suggesting that most studies were conducted using younger adult mice. Seven percent of studies did not report the age of the mice. On average, sedentary control and exercise training groups included 9 ± 9 (mean ± SD) mice and 10 ± 8 mice per study, respectively.

Training Protocols

There was a wide range of treadmill training protocols reported. Most studies included information about the training protocol components: frequency of exercise (days/week), velocity of the treadmill, incline of the treadmill, length of each session (time in minutes), and the duration of the exercise training (weeks). Treadmill velocity was reported as m/min, m/s, or cm/s. Treadmill incline was reported in degrees or % incline. Velocity and incline were converted to m/min and degrees for data analysis. The mode for each parameter was: frequency of 5 days/week (91% of studies, range: 2–7 days/week), a treadmill velocity of 15 m/min (19%, 5.25–33.8 m/min), 10° of treadmill incline (38%, 0–25°), 60 min/session (64%, 9–540 min/session), and a duration of 4 weeks (53%, 2–16 weeks). Fifty studies (48%) included information for all components of the training protocol. The number of studies with missing exercise protocol data was: frequency: $n = 3$ missing, treadmill velocity: $n = 38$, treadmill incline: $n = 39$, time per session: $n = 4$, and training duration: $n = 2$. Studies not reporting some or all these components typically listed exercise intensity instead. Exercise intensities were reported as low, moderate, high or as a percentage of maximum.

Exercise Tests

Performance outcomes (i.e., time, work, or distance) based on the results of an exercise performance test were reported in 78 studies. Exercise testing was not uniform in these studies

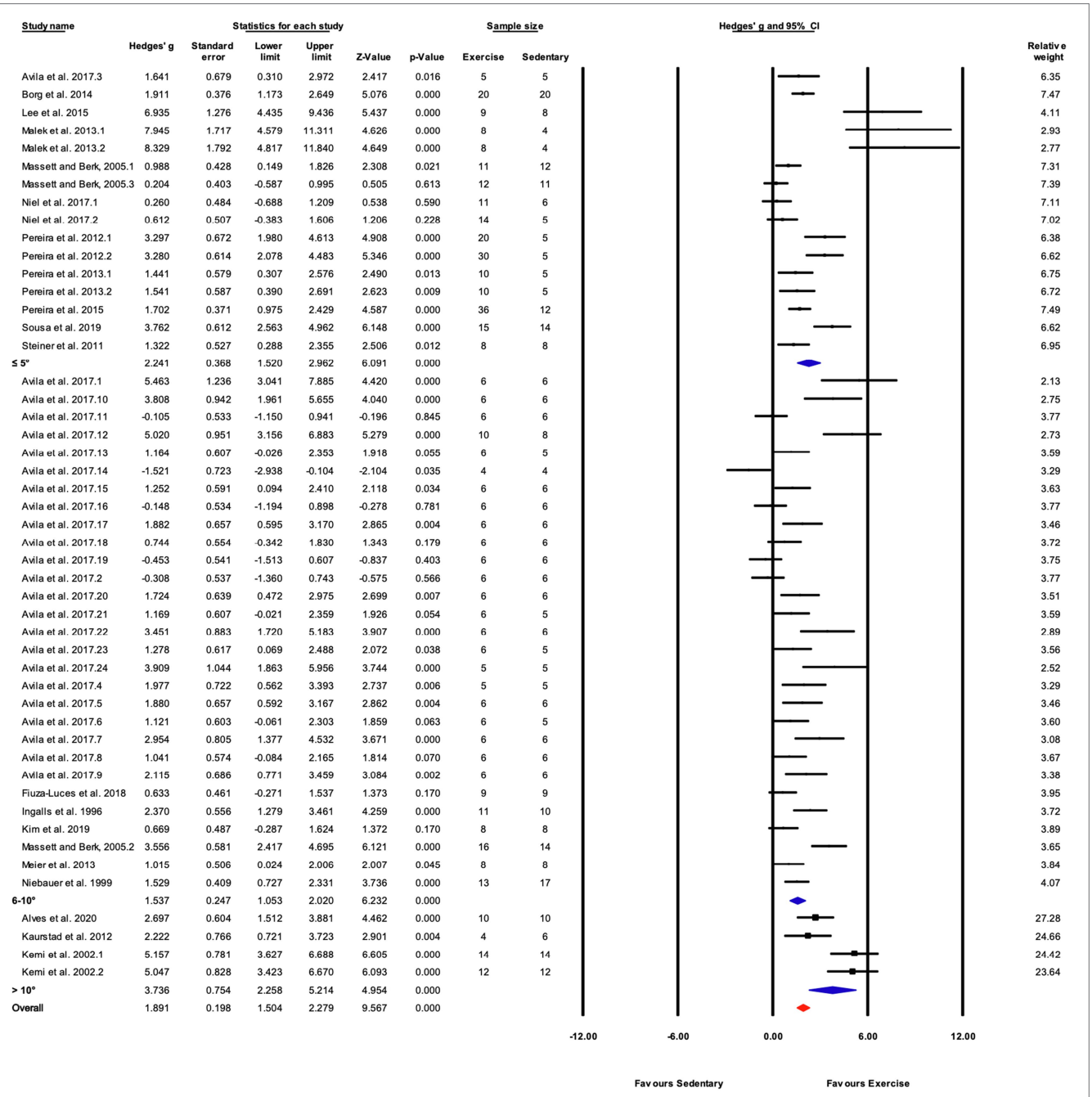


FIGURE 6 | Forest plot of the between-group comparisons of the effect of training protocol duration on performance-based markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

and therefore was categorized by the testing protocol or outcome. Most testing protocols fell into three primary categories: Graded Exercise Testing (GXT), Incremental Load Testing (ILT), and $\dot{V}O_{2max}$. A GXT that included incremental increases in both treadmill velocity and incline was utilized in 46% of studies reporting performance-based outcomes. ILT, an incremental increase in treadmill velocity and no change in incline, was reported in 46% of studies, and a $\dot{V}O_{2max}$ protocol, measuring maximal oxygen consumption during exercise, was used in 8% of studies reporting performance-based outcomes. Submaximal endurance tests at a constant workload were used in a few other studies (Steiner et al., 2011; Kim et al., 2019).

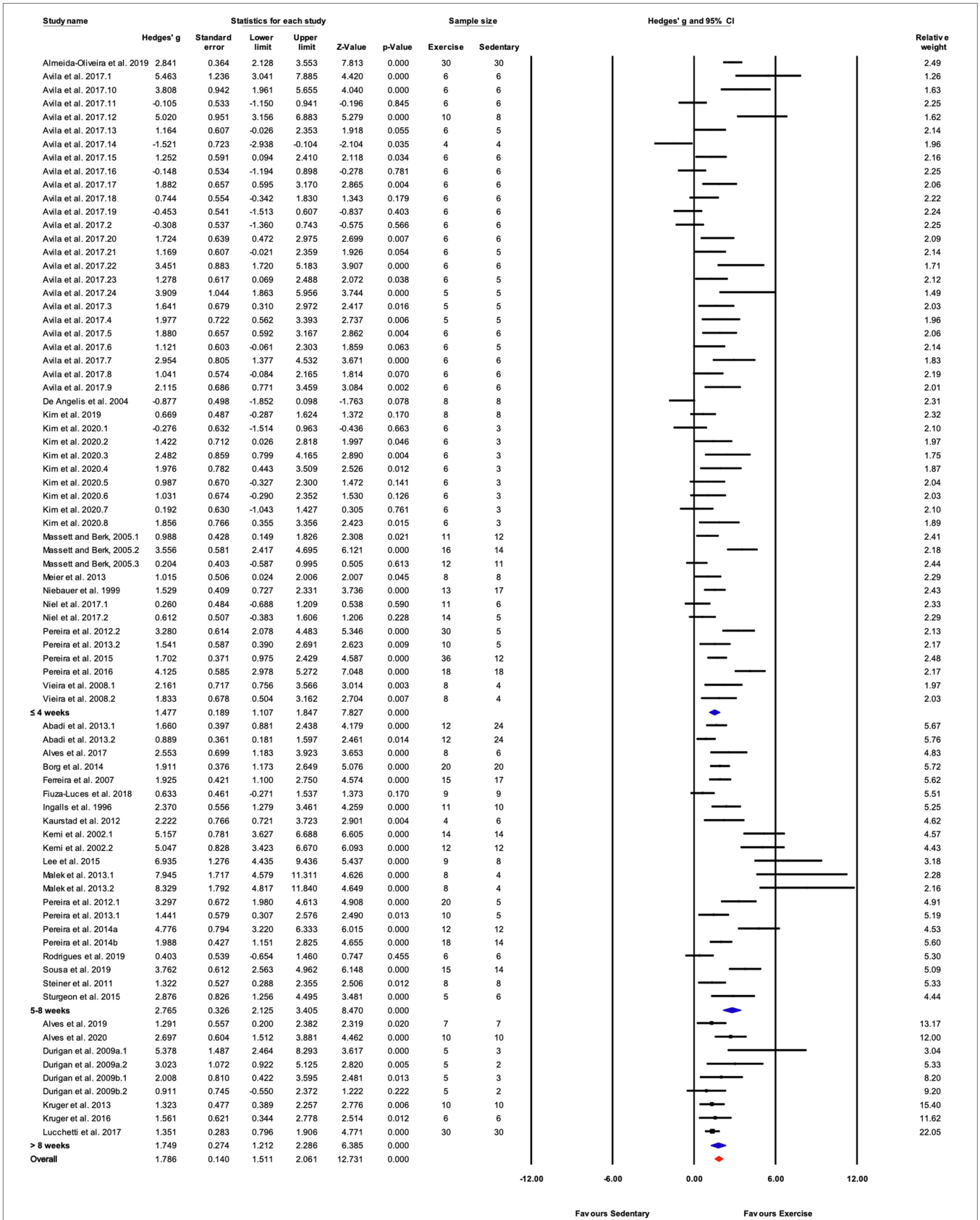


FIGURE 7 | Continued

FIGURE 7 | Forest plot of the between-group comparisons of the effect of treadmill incline on performance-based markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

Outcomes

Most studies (74%) used a measure of exercise performance as a marker of exercise training efficacy (**Table 1**). Performance was measured during an exercise test (above) and reported as time, distance, maximal speed/velocity, or work. Differences in these outcome measures were compared between sedentary and exercise training groups. Twenty studies (19%) assessed training efficacy using biochemical measures including citrate synthase or succinate dehydrogenase enzyme activity in skeletal muscle, blood lactate concentrations, or mitochondrial DNA copy number (**Table 1**). Other outcome measures used were heart weight or heart weight to body weight ratios (Suominen et al., 1980; Foryst-Ludwig et al., 2011; Han, 2013), the number of type 2 skeletal muscle fibers (Wernig et al., 1991), and the systolic blood pressure before and after training (Pinto et al., 2015).

Meta-Analysis

Overall Effect Size and Heterogeneity

The data from 105 studies was aggregated in the random effect model for the meta-analysis (**Figure 4**). The overall effect of exercise was statistically significant, with high heterogeneity (Hedges's $g = 1.70$, 95% CI = 1.47–1.94, $p < 0.05$, $Tau^2 = 1.14$, $I^2 = 80.4\%$, prediction interval = −0.43–3.84). To investigate the heterogeneity across studies, subgroup analysis was performed using 10 moderator variables: strain, age, sex, training intensity, velocity, incline, time/session, duration, performance test, and the type of outcome variable (e.g., performance-based, biochemical, etc.). Performance tests included GXT, ILT, maximal oxygen consumption test ($\dot{V}O_{2max}$). **Table 2** shows the outcome of the subgroup analysis for each of the moderators. Treadmill incline, training duration, exercise performance test protocol, and outcome variable showed significant differences between subgroups. We also performed meta-regression to determine the percentage of heterogeneity explained by each moderator subgroup and by the combination of moderators related to the exercise training protocol. The results for the individual moderators are shown in **Table 2**. Five moderators, treadmill velocity, treadmill incline, exercise session time, performance test, and outcome variable category, each showed significant associations between moderator value and exercise training response. When training frequency, treadmill velocity and incline, time per exercise session, and training duration were included in the meta-regression as continuous variables, this model accounted for 0% of the between-study variance, suggesting that other factors are also contributing to differences between studies.

Based on the significant difference observed for subgroups of outcome variables (e.g., performance-based vs. biochemical; **Table 2**), separate meta-analyses were performed for studies with performance-based outcome variables and studies which reported biochemical-related outcome variables. There were too few studies coded as "Other" to support a separate analysis of studies in that category. Thus, two separate meta-analyses were done on two different groups of studies: (1) a group of studies with performance-based outcome variables, and (2) another group of studies that reported biochemical-related outcome variables.

Results for Performance-Based Outcome

Seventy-eight (78) studies out of 105 (74%) included in the meta-analysis assessed performance-based outcome variables such as exhaustion time, maximum velocity, or work. The overall effect of exercise training on performance-based outcome variables from those studies was significant, with high between-study heterogeneity (Hedges' $g = 1.85$, 95% CI = 1.55–2.15, $p < 0.05$, Q-value = 390.13, df = 77, $Tau^2 = 1.35$, $I^2 = 80.3\%$, prediction interval = −0.48 to 4.18). A summary of the subgroup analysis performed to investigate the heterogeneity across the studies reporting performance-based outcome variables is shown in **Figure 5**. Significant differences between subgroups were observed for treadmill incline, training duration, and the type of exercise test. Non-significant results were obtained from the subgroup analyses of strain, age, sex, exercise intensity, treadmill velocity, and time/session. Results for moderator variables with significant differences between subgroups are described below.

Grouped by Treadmill Incline

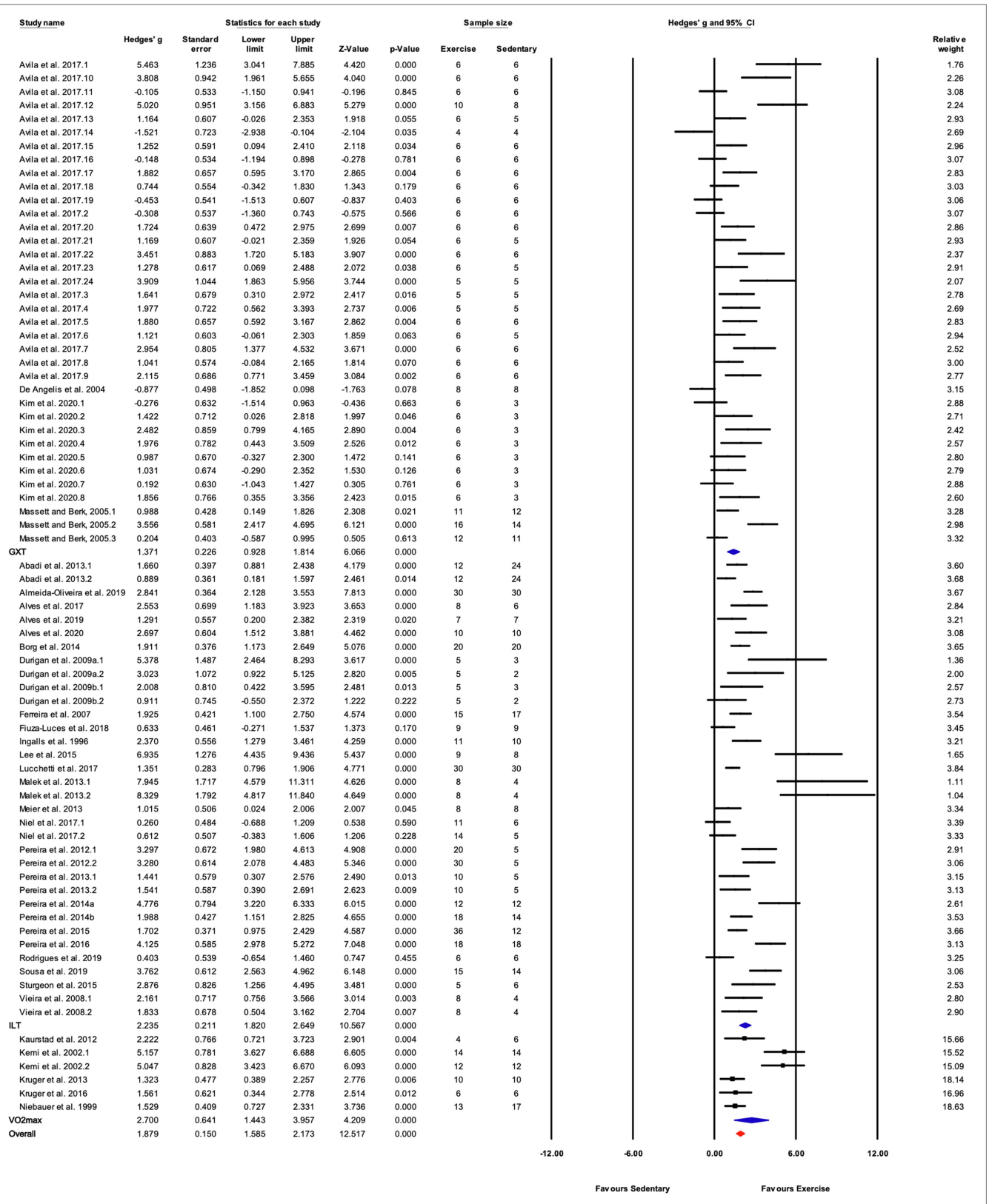
When studies were divided based on treadmill incline, significant differences between trained and sedentary groups were observed regardless of the incline (**Figures 5, 6**). Studies that incorporated an incline >10° had a greater response to training relative to those with inclinations of ≤5° and 6–10° (Q-between = 8.96, df = 2, $p < 0.05$, $I^2 = 82.6\%$).

Grouped by Training Duration

Studies were divided into three ranges of training duration: "≤4," "5–8," and ">8 weeks." Each training duration was associated with a significant increase in performance (**Figures 5**, **7**). Mice training for 5–8 weeks had a greater response than those training for a shorter duration ("≤4 weeks"), or a longer duration (">8 weeks"; Q-between = 11.69, df = 2, $p < 0.05$, $I^2 = 80.3\%$).

Grouped by Exercise Test

There were three subgroups in the covariate exercise test: "GXT," "ILT," and "$\dot{V}O_{2max}$." All tests were associated with significant increases in training responses. A significant difference was observed between testing protocols (Q-between = 9.52, df = 2,



Study name	Hedges' g	Standard error	Lower limit	Upper limit	Z-Value	p-Value	Exercise	Sedentary	Relative weight
Avila et al. 2017.1	5.463	1.236	3.041	7.885	4.420	0.000	6	6	1.76
Avila et al. 2017.10	3.808	0.942	1.961	5.655	4.040	0.000	6	6	2.26
Avila et al. 2017.11	-0.105	0.533	-1.150	0.941	-0.196	0.845	6	6	3.08
Avila et al. 2017.12	5.020	0.951	3.156	6.883	5.279	0.000	10	8	2.24
Avila et al. 2017.13	1.164	0.607	-0.026	2.353	1.918	0.055	6	5	2.93
Avila et al. 2017.14	-1.521	0.723	-2.938	-0.104	-2.104	0.035	4	4	2.69
Avila et al. 2017.15	1.252	0.591	0.094	2.410	2.118	0.034	6	6	2.96
Avila et al. 2017.16	-0.148	0.534	-1.194	0.898	-0.278	0.781	6	6	3.07
Avila et al. 2017.17	1.882	0.657	0.595	3.170	2.865	0.004	6	6	2.83
Avila et al. 2017.18	0.744	0.554	-0.342	1.830	1.343	0.179	6	6	3.03
Avila et al. 2017.19	-0.453	0.541	-1.513	0.607	-0.837	0.403	6	6	3.06
Avila et al. 2017.2	-0.308	0.537	-1.360	0.743	-0.575	0.566	6	6	3.07
Avila et al. 2017.20	1.724	0.639	0.472	2.975	2.699	0.007	6	6	2.86
Avila et al. 2017.21	1.169	0.607	-0.021	2.359	1.926	0.054	6	5	2.93
Avila et al. 2017.22	3.451	0.883	1.720	5.183	3.907	0.000	6	6	2.37
Avila et al. 2017.23	1.278	0.617	0.069	2.488	2.072	0.038	6	5	2.91
Avila et al. 2017.24	3.909	1.044	1.863	5.956	3.744	0.000	5	5	2.07
Avila et al. 2017.3	1.641	0.679	0.310	2.972	2.417	0.016	5	5	2.78
Avila et al. 2017.4	1.977	0.722	0.562	3.393	2.737	0.006	5	5	2.69
Avila et al. 2017.5	1.880	0.657	0.592	3.167	2.862	0.004	6	6	2.83
Avila et al. 2017.6	1.121	0.603	-0.061	2.303	1.859	0.063	6	5	2.94
Avila et al. 2017.7	2.954	0.805	1.377	4.532	3.671	0.000	6	6	2.52
Avila et al. 2017.8	1.041	0.574	-0.084	2.165	1.814	0.070	6	6	3.00
Avila et al. 2017.9	2.115	0.686	0.771	3.459	3.084	0.002	6	6	2.77
De Angelis et al. 2004	-0.877	0.498	-1.852	0.098	-1.763	0.078	8	8	3.15
Kim et al. 2020.1	-0.276	0.632	-1.514	0.963	-0.436	0.663	6	3	2.88
Kim et al. 2020.2	1.422	0.712	0.026	2.818	1.997	0.046	6	3	2.71
Kim et al. 2020.3	2.482	0.859	0.799	4.165	2.890	0.004	6	3	2.42
Kim et al. 2020.4	1.976	0.782	0.443	3.509	2.526	0.012	6	3	2.57
Kim et al. 2020.5	0.987	0.670	-0.327	2.300	1.472	0.141	6	3	2.80
Kim et al. 2020.6	1.031	0.674	-0.290	2.352	1.530	0.126	6	3	2.79
Kim et al. 2020.7	0.192	0.630	-1.043	1.427	0.305	0.761	6	3	2.88
Kim et al. 2020.8	1.856	0.766	0.355	3.356	2.423	0.015	6	3	2.60
Massett and Berk, 2005.1	0.988	0.428	0.149	1.826	2.308	0.021	11	12	3.28
Massett and Berk, 2005.2	3.556	0.581	2.417	4.695	6.121	0.000	16	14	2.98
Massett and Berk, 2005.3	0.204	0.403	-0.587	0.995	0.505	0.613	12	11	3.32
GXT	1.371	0.226	0.928	1.814	6.066	0.000			
Abadi et al. 2013.1	1.660	0.397	0.881	2.438	4.179	0.000	12	24	3.60
Abadi et al. 2013.2	0.889	0.361	0.181	1.597	2.461	0.014	12	24	3.68
Almeida-Oliveira et al. 2019	2.841	0.364	2.128	3.553	7.813	0.000	30	30	3.67
Alves et al. 2017	2.553	0.699	1.183	3.923	3.653	0.000	8	6	2.84
Alves et al. 2019	1.291	0.557	0.200	2.382	2.319	0.020	7	7	3.21
Alves et al. 2020	2.697	0.604	1.512	3.881	4.462	0.000	10	10	3.08
Borg et al. 2014	1.911	0.376	1.173	2.649	5.076	0.000	20	20	3.65
Durigan et al. 2009a.1	5.378	1.487	2.464	8.293	3.617	0.000	5	3	1.36
Durigan et al. 2009a.2	3.023	1.072	0.922	5.125	2.820	0.005	5	2	2.00
Durigan et al. 2009b.1	2.008	0.810	0.422	3.595	2.481	0.013	5	3	2.57
Durigan et al. 2009b.2	0.911	0.745	-0.550	2.372	1.222	0.222	5	2	2.73
Ferreira et al. 2007	1.925	0.421	1.100	2.750	4.574	0.000	15	17	3.54
Fiuza-Luces et al. 2018	0.633	0.461	-0.271	1.537	1.373	0.170	9	9	3.45
Ingalls et al. 1996	2.370	0.556	1.279	3.461	4.259	0.000	11	10	3.21
Lee et al. 2015	6.935	1.276	4.435	9.436	5.437	0.000	9	8	1.65
Lucchetti et al. 2017	1.351	0.283	0.796	1.906	4.771	0.000	30	30	3.84
Malek et al. 2013.1	7.945	1.717	4.579	11.311	4.626	0.000	8	4	1.11
Malek et al. 2013.2	8.329	1.792	4.817	11.840	4.649	0.000	8	4	1.04
Meier et al. 2013	1.015	0.506	0.024	2.006	2.007	0.045	8	8	3.34
Niel et al. 2017.1	0.260	0.484	-0.688	1.209	0.538	0.590	11	6	3.39
Niel et al. 2017.2	0.612	0.507	-0.383	1.606	1.206	0.228	14	5	3.33
Pereira et al. 2012.1	3.297	0.672	1.980	4.613	4.908	0.000	20	5	2.91
Pereira et al. 2012.2	3.280	0.614	2.078	4.483	5.346	0.000	30	5	3.06
Pereira et al. 2013.1	1.441	0.579	0.307	2.576	2.490	0.013	10	5	3.15
Pereira et al. 2013.2	1.541	0.587	0.390	2.691	2.623	0.009	10	5	3.13
Pereira et al. 2014a	4.776	0.794	3.220	6.333	6.015	0.000	12	12	2.61
Pereira et al. 2014b	1.988	0.427	1.151	2.825	4.655	0.000	18	14	3.53
Pereira et al. 2015	1.702	0.371	0.975	2.429	4.587	0.000	36	12	3.66
Pereira et al. 2016	4.125	0.585	2.978	5.272	7.048	0.000	18	18	3.13
Rodrigues et al. 2019	0.403	0.539	-0.654	1.460	0.747	0.455	6	6	3.25
Sousa et al. 2019	3.762	0.612	2.563	4.962	6.148	0.000	15	14	3.06
Sturgeon et al. 2015	2.876	0.826	1.256	4.495	3.481	0.000	5	6	2.53
Vieira et al. 2008.1	2.161	0.717	0.756	3.566	3.014	0.003	8	4	2.80
Vieira et al. 2008.2	1.833	0.678	0.504	3.162	2.704	0.007	8	4	2.90
ILT	2.235	0.211	1.820	2.649	10.567	0.000			
Kaurstad et al. 2012	2.222	0.766	0.721	3.723	2.901	0.004	4	6	15.66
Kemi et al. 2002.1	5.157	0.781	3.627	6.688	6.605	0.000	14	14	15.52
Kemi et al. 2002.2	5.047	0.828	3.423	6.670	6.093	0.000	12	12	15.09
Kruger et al. 2013	1.323	0.477	0.389	2.257	2.776	0.006	10	10	18.14
Kruger et al. 2016	1.561	0.621	0.344	2.778	2.514	0.012	6	6	16.96
Niebauer et al. 1999	1.529	0.409	0.727	2.331	3.736	0.000	13	17	18.63
VO2max	2.700	0.641	1.443	3.957	4.209	0.000			
Overall	1.879	0.150	1.585	2.173	12.517	0.000			

FIGURE 8 | Forest plot of the between-group comparisons of the effect of exercise performance test protocol on performance-based markers of training efficacy. GXT, graded exercise test; ILT, incremental load test; $\dot{V}O_{2max}$, maximal oxygen consumption test. Standardized mean differences were calculated as Hedges' *g*.

(Continued)

FIGURE 8 | Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

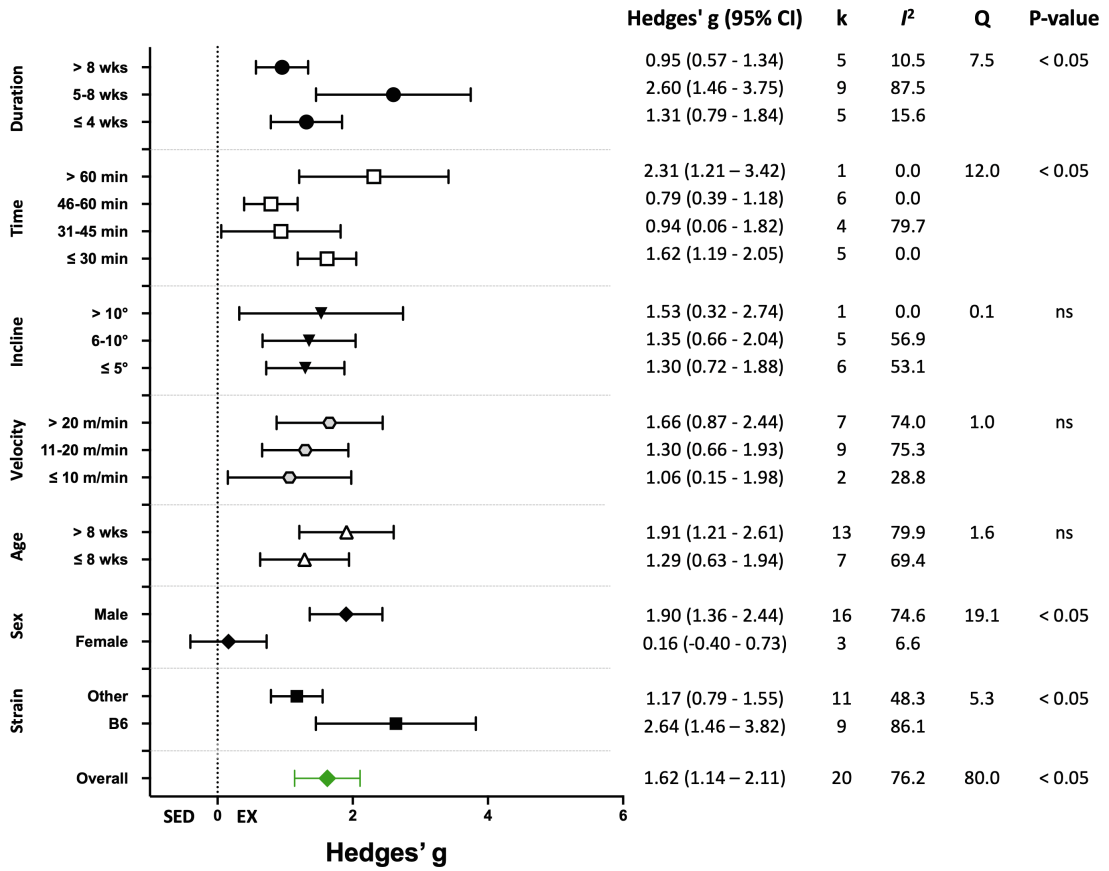


FIGURE 9 | Adjusted effect sizes of the between-group comparisons for the predefined moderators on biochemical exercise training outcomes. GXT, graded exercise test; ILT, incremental load test; $\dot{V}O_{2max}$, maximal oxygen consumption test; B6, C57BL/6 mice; k, number of studies in each subgroup; I^2, measure of heterogeneity; *Q*, Cochran's *Q*; *p*, value of *p* for heterogeneity analysis (overall) or differences between subgroups; ns, non-significant $p > 0.05$.

$p < 0.05$, $I^2 = 80.6\%$; **Figures 5**, **8**). The largest effect of training was observed for studies utilizing the $\dot{V}O_{2max}$ test, followed by ILT, and GXT; however, the 95% CI for the $\dot{V}O_{2max}$ group included the point estimate of the ILT subgroup. ILT was significantly greater than GXT (**Figures 5**, **8**).

A multivariate meta-regression that included training frequency (day/week), treadmill velocity (m/min) and incline (degrees), time/session (min), and training duration (weeks) was performed to determine the association between exercise training components and performance outcomes. Thirty-five studies were included in the meta-regression. None of the coefficients in a multivariate meta-regression were significant and overall, the model did not explain any of the between-study variance in effect size ($R^2 = 0.0$).

Results for Biochemical Outcomes

Nineteen percent of the studies (20 of 105) reported biochemical outcomes, including citrate synthase or succinate dehydrogenase activity, or mitochondrial DNA copy number, or lactate levels as the indicators of training efficacy. The overall effect of exercise training on biochemical-based outcome variables was significant, with high heterogeneity (Hedges' $g = 1.62$, 95% CI = 1.14–2.11, $p < 0.05$, *Q*-value = 80.0, df = 19, $Tau^2 = 0.84$, $I^2 = 76.2\%$, prediction interval = −0.37–3.62). A summary of the analyses for the moderator variables analyzed is shown in **Figure 9**. Significant improvements in biochemical outcomes with exercise training were shown in male mice and in studies with a training duration of 5–8 weeks. Significant subgroup differences also were observed for mouse strain and time/session (**Figure 9**). Results for moderator variables with significant differences between subgroups are described below.

Grouped by Mouse Strain

Studies were divided into two mouse strain subcategories, C57BL/6 and "Other." The "Other" category included six mouse strains and accounted for 11 of 20 studies (55%). Both cohorts

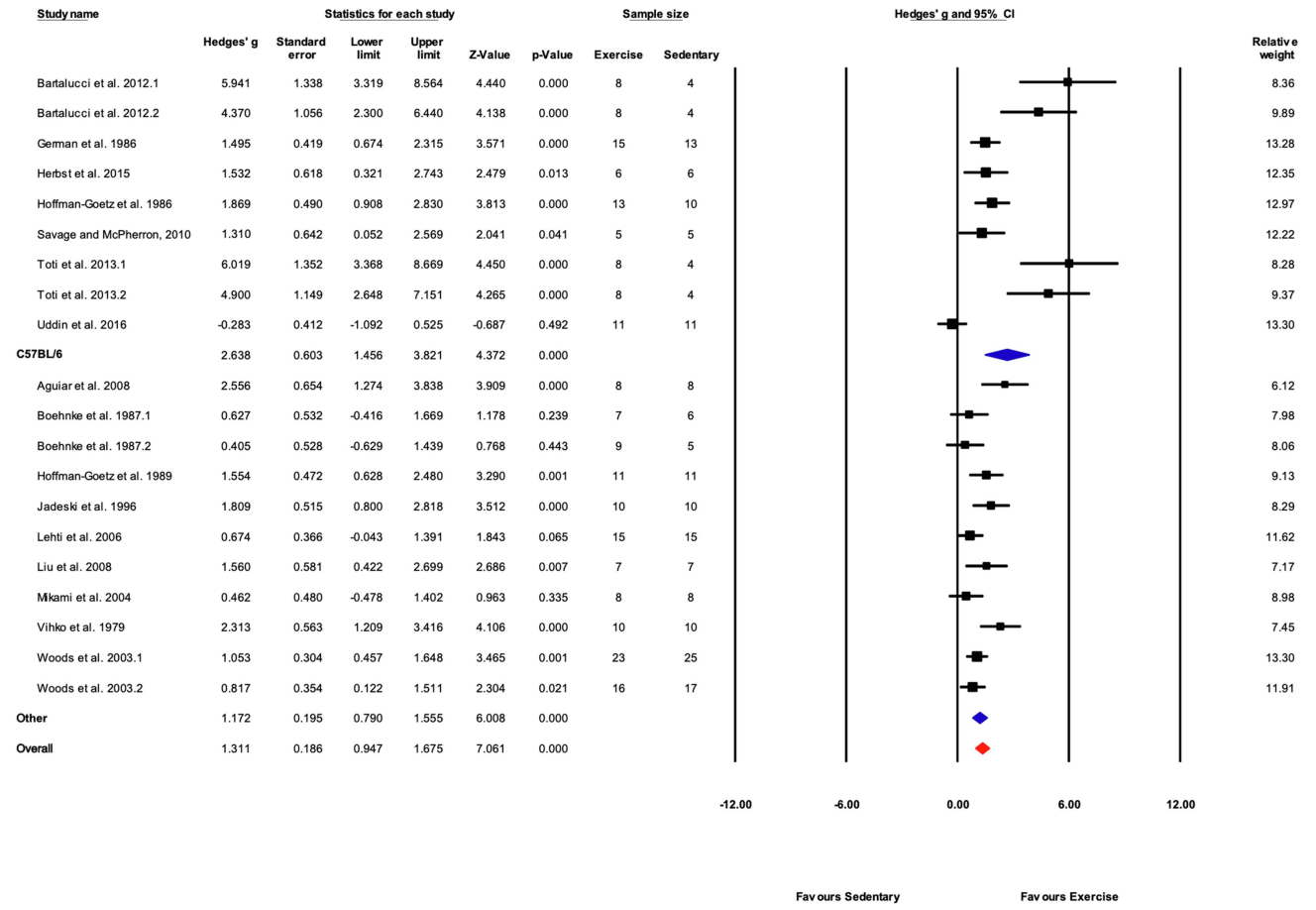


FIGURE 10 | Forest plot of the between-group comparisons of the effect of mouse strain on biochemical trait markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

showed significant responses to training. The response to training was significantly greater in C57BL/6 mice compared with other strains (Q-between = 5.34, df = 1, $p < 0.05$, $I^2 = 76.2\%$; **Figures 9, 10**).

Grouped by Sex

Subgroup analysis shows a significant difference between male and female mice (Q-between = 19.1, df = 1, $p < 0.05$, $I^2 = 77.1\%$). Only three studies included female mice compared with 16 using male mice. Female mice showed a non-significant response to training ($p > 0.05$) compared with sedentary controls (**Figures 9**, **11**).

Grouped by Time

Exercise time per session was divided into four subgroups: "≤30," "31–45," "46–60," and ">60 min" consisting of 5, 4, 6, and 1 studies, respectively. Exercise training elicited a significant effect in each exercise time subgroup. Significant differences between subgroups were present (Q-between = 11.99, df = 3, $p < 0.05$, $I^2 = 56.6\%$). The largest effect was in the one study in the ">60 min" subcategory, followed by the "≤30 min" subcategory (**Figures 9**, **12**).

Grouped by Training Duration

As in the overall and performance-based outcome analyses, the effect of exercise training was significant for all training durations (**Figures 9**, **13**). The response to training was significantly greater in the "5–8 weeks" group compared with "≤4" and ">8 weeks" (Q-between = 7.48, df = 2, $p < 0.05$, $I^2 = 76.4\%$).

A multivariate meta-regression that included training frequency (day/week), treadmill velocity (m/min) and incline (degrees), time/session (min), and training duration (weeks) was performed to determine the association between exercise training components and biochemical outcomes. Eleven studies had complete data for each variable and were included in the meta-regression analysis. Although none of the coefficients in the model were significant, 100% of the between-study variance was explained by the model ($R^2 = 1.00$).

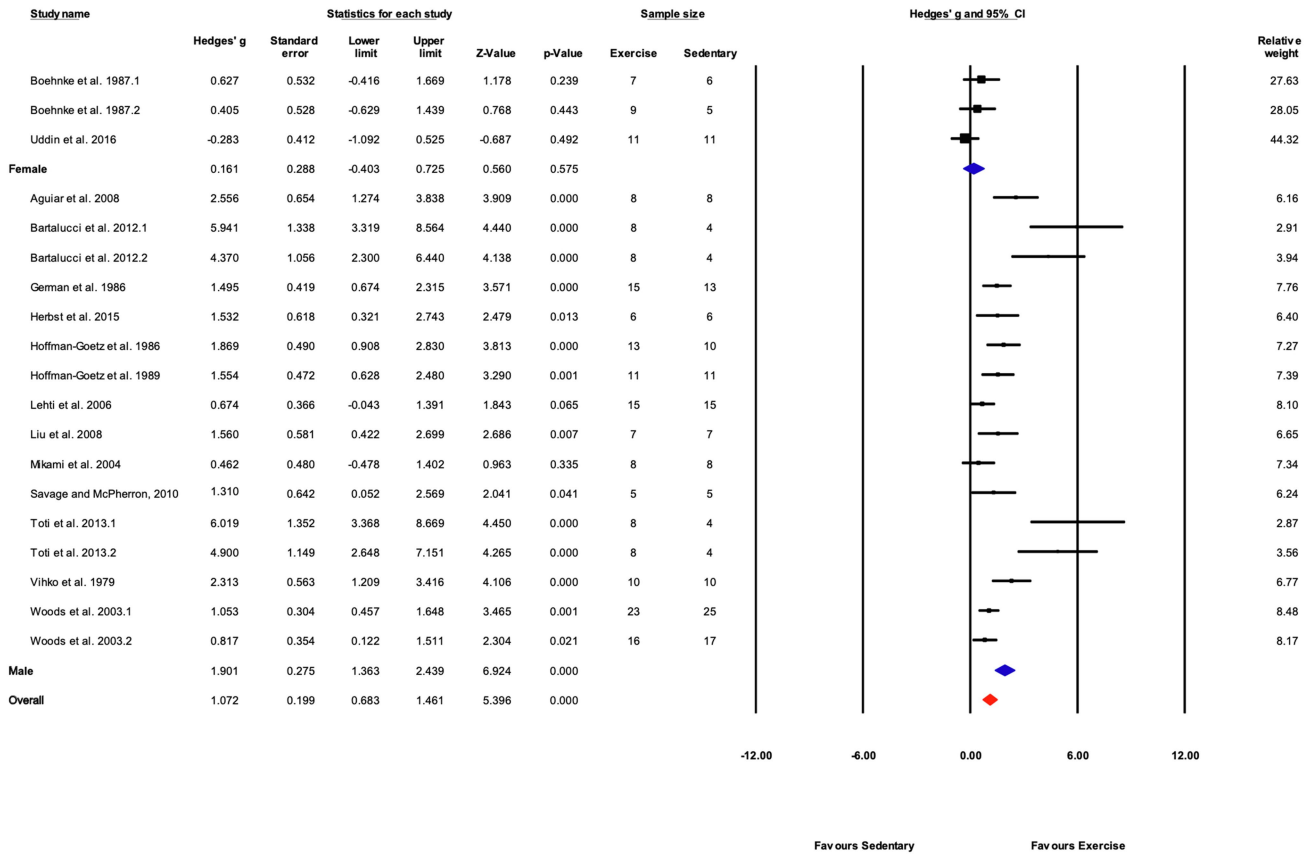


FIGURE 11 | Forest plot of the between-group comparisons of the effect of sex on biochemical trait markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

DISCUSSION

The main findings of this systematic review and meta-analysis of mouse exercise training studies are: (1) a relatively small number of studies incorporating exercise training report a "classical" measure of training efficacy; (2) many studies do not report complete information regarding the exercise training protocol; (3) the majority of exercise training studies utilize male mice only; (4) exercise training significantly increases measures of training efficacy; and (5) exercise prescription parameters do not explain a significant amount of variation between studies when changes in exercise performance are used as a marker for training efficacy.

Our systematic review identified 164 full-text articles that included a treadmill training protocol with untreated mice assigned to either a sedentary control group or exercise training group. Of these, approximately 35% included a "classical" marker of training efficacy. Increases in skeletal muscle enzyme activity, mitochondrial DNA, and/or changes in skeletal muscle fiber types are possible markers for adaptations to endurance exercise training (Booth et al., 2010). An increase in peak or maximal oxygen consumption is often considered the gold standard in human-based endurance exercise training studies. In animal studies, changes in exercise performance are typically used as a surrogate for maximal oxygen consumption (Fuller and Thyfault, 2021). Therefore, only studies including these or other well-known markers for exercise training adaptations were included (Holloszy and Coyle, 1984; Hellsten and Nyberg, 2015). The majority of studies that were excluded for lack of such a marker utilized body weight differences between sedentary and exercise-trained groups as a general marker for exercise training. Although lower body weights in the exercise training group might be related to increased physical activity, body weight differences alone do not necessarily indicate that the exercise training elicited beneficial biochemical and/or cardiorespiratory fitness adaptations. For purposes of replication and thorough analysis of the responses to exercise, exercise training studies should include all relevant information regarding the training protocol such as frequency, intensity, and duration (Booth et al., 2010). All protocol information was included in 48% of the studies. Treadmill velocity (38%) and incline (39%) were the most frequently omitted variables. Most reported

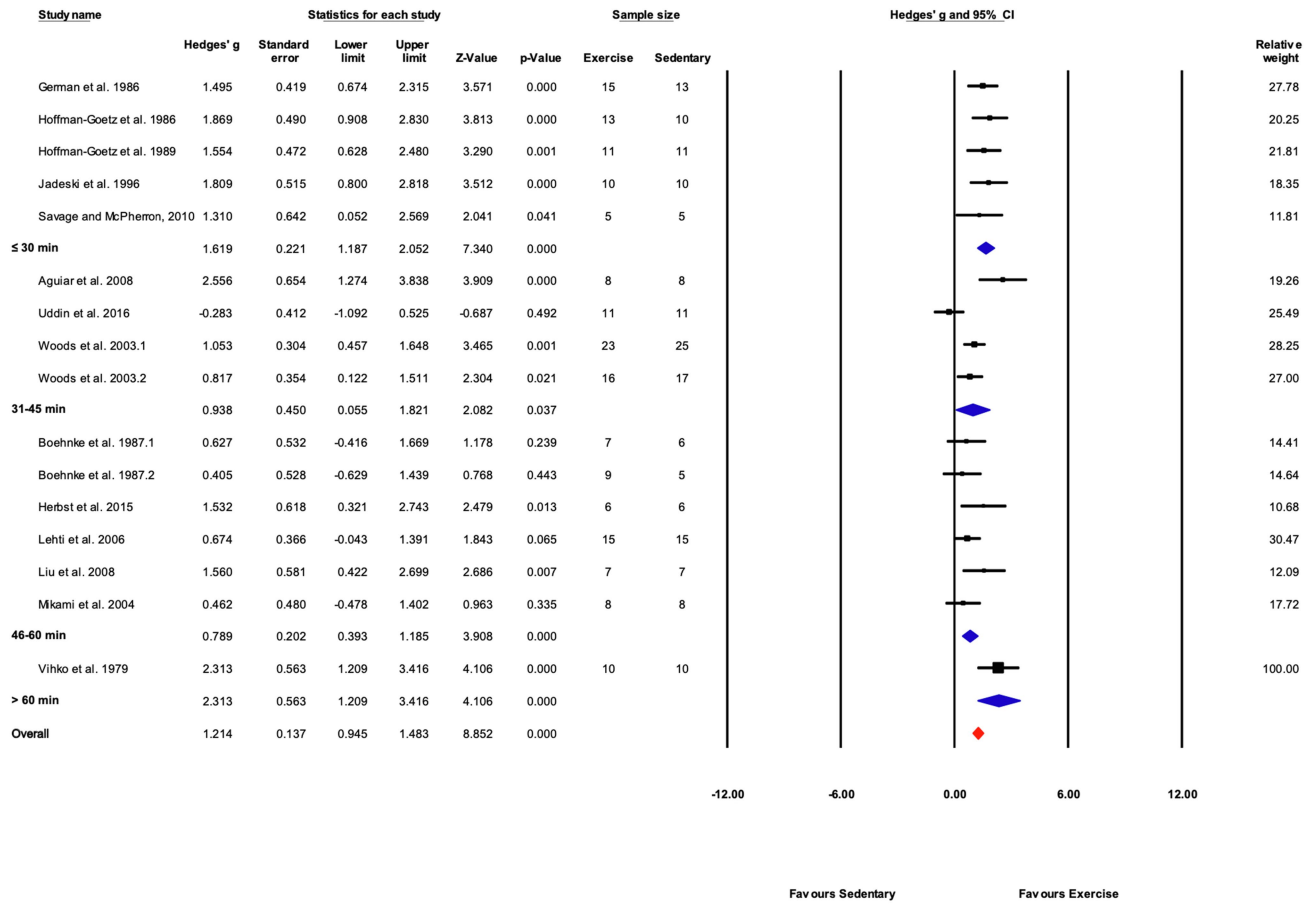


FIGURE 12 | Forest plot of the between-group comparisons of the effect of exercise training session time on biochemical trait markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

frequency, session time and duration. Exercise intensity was reported in 68% of studies, but the basis for qualifiers low, moderate, and high were unclear. Treadmill velocity and incline were frequently omitted when exercise intensity as a percentage of maximum was reported. Collectively, these results indicate that treadmill-based exercise training studies in mice frequently do not report all the components of the exercise training program or well-accepted adaptations to exercise training as indicators of training efficacy.

Mouse strain, sex, and age have been reported to influence exercise training responses. Overall, these moderators had limited effects on exercise training responses. When outcome variables were divided into performance-based and biochemical outcomes, sex and mouse strain significantly influenced biochemical responses to training (**Figure 9**). Male mice had significantly greater biochemical adaptations to exercise training than female mice. In contrast, performance-based outcomes were somewhat greater in females than males, but not significantly so (**Figure 5**). In a direct comparison, Kemi et al. (2002) reported that $\dot{V}O_{2max}$ was significantly greater in trained female mice than in similarly trained male mice. Similarly, exercise training-induced cardiac hypertrophy was greater in female mice compared with males (Foryst-Ludwig et al., 2011). However, less than 10% of the included studies utilized female mice and only three full-text articles included both male and female mice (Kemi et al., 2002; Foryst-Ludwig et al., 2011; Abadi et al., 2013). Therefore, additional studies are needed investigating the responses to endurance exercise training in female mice as well as studies directly comparing responses in mice of both sexes.

The influence of mouse strain was not significant overall (**Table 2**), but was significant in studies measuring biochemical markers of exercise training. For subgroup analyses, strains were coded as C57BL/6 or "Other." The "Other" group included data from 27 strains. As with sex comparisons, only three articles included data from multiple mouse strains (Massett and Berk, 2005; Avila et al., 2017; Kim et al., 2020). Each of those publications reported significant strain-dependent changes in exercise capacity in response to exercise training. However, those findings were not supported by the results of the current

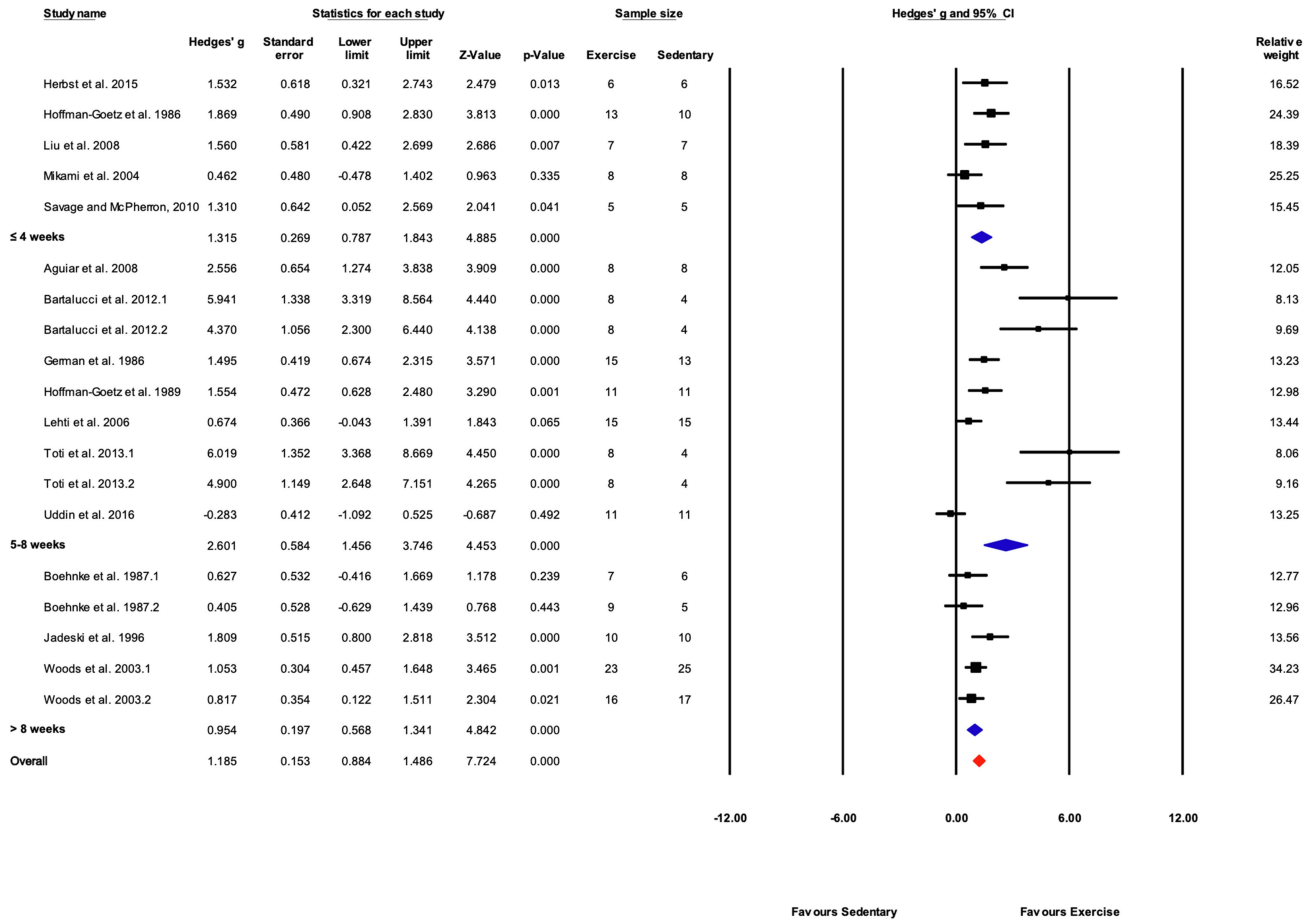


FIGURE 13 | Forest plot of the between-group comparisons of the effect of training protocol duration on biochemical trait markers of training efficacy. Standardized mean differences were calculated as Hedges' *g*. Overall analysis was conducted using a random effects model. Values to the left of zero (Favors Sedentary) indicates the sedentary group had a greater response. Values to the right of zero (Favors Exercise) indicates a greater response in the exercise training group. The size of the black squares indicates the weight of the study-specific estimates. Blue diamond indicates pooled estimate of random effects model for each subgroup. Red diamond indicates overall pooled estimate of random effects model.

study for performance-based outcomes. One possible explanation for this disparity is the "Other" strain category is composed of too many individual strains, leading to a high level of variation across subject populations and training protocols. However, the precision and dispersion of the effect estimates are similar for both C57BL/6 and "Other" subgroups suggesting that the variability in response to training is comparable. Thus, the strain-dependent differences in changes in exercise capacity with exercise training reported by Massett and colleagues (Massett and Berk, 2005; Avila et al., 2017; Kim et al., 2020) might be specific to the exercise training and testing paradigm used in those studies. Each of those studies utilized similar exercise training parameters with some strain-specific adjustments which facilitated direct comparisons with minimal variation between training protocols. Conversely, responses in C57BL/6 mice were significantly greater than other strains for biochemical markers of exercise training. This result implies that C57BL/6 mice show greater biochemical adaptations to exercise training than mice from other strains. This contrasts with performance-based outcomes where C57BL/6 mice have low to moderate responses to training compared with other strains (Massett and Berk, 2005; Avila et al., 2017; Kim et al., 2020). Future research investigating the effect of mouse strain on exercise training responses should consider including multiple strains within the same study design and measuring both performance-based and biochemical markers of training efficacy. Collectively, the findings regarding the contribution of sex and mouse strain on responses to exercise training suggest that direct comparisons within a given experimental design might yield results different from those obtained in a pooled analysis of the published studies utilizing individual mouse strains.

Overall, exercise training elicited significant increases in exercise training-associated outcomes. Heterogeneity was high for the combined analysis as well as for outcome-specific analyses. Therefore, subgroup analyses were performed for the combined data and for performance and biochemical outcomes separately. In the combined analysis, the greatest percentage of variation in the effect size was explained by exercise test

protocol subgroups (**Table 2**). Exercise training parameters of frequency, treadmill velocity and incline, exercise session time, and training duration also were investigated to determine their contribution to the heterogeneity between studies/as potential moderator variables. In the overall analysis of 105 studies and in the separate analyses based on training outcome, subgroup analysis was significant for training duration (**Table 2; Figures 5**, **9**). Studies utilizing a training duration of 5–8 weeks had significantly greater outcomes than those incorporating longer or shorter periods. Typically, exercise training protocols include one or more weeks during which training time/intensity is progressively increased until the target parameters are reached. For studies ≤4 weeks, the target workload might be sustained for too short a period (e.g., 2 weeks) to elicit maximal training responses. Protocols longer than 8 weeks showed effects comparable to those ≤4 weeks, implying that longer duration training protocols might hinder adaptations to training. The mechanism for this is unclear, but declines in performance with prolonged training, especially at higher intensities, can be associated with overtraining syndrome (Pereira et al., 2012; Meeusen et al., 2013). These results suggest that regardless of training outcome a training duration of 5–8 weeks appears optimal for adaptations to treadmill training in mice.

Significant subgroup differences for treadmill incline were observed in the combined analysis and for performance-based outcomes (**Table 2; Figures 5, 6**). Exercise training protocols utilizing an incline above 10° had the largest effect on training outcomes in the combined and performance-based analyses (**Table 2; Figures 5, 6**). The higher incline should require more work/greater effort and therefore, might elicit greater adaptations to training (Hoydal et al., 2007; Poole et al., 2020). Kemi and colleagues (Kemi et al., 2002; Hoydal et al., 2007) reported that the best estimates of mouse $\dot{V}O_{2max}$ are obtained during treadmill exercise at inclines between 15° and 35°. They demonstrated significant improvements in maximal oxygen consumption after 8 weeks of training at 25° in male and female C57BL/6J mice (Kemi et al., 2002). Therefore, they recommended an inclination of 25° as optimal for exercise training. However, Petrosino et al. (2016) limited the treadmill incline to 15° in the development of their exercise testing protocol because they observed that mice had difficulty maintaining gait at inclinations above 15°. Although gait changes during treadmill running in rodents, including raising of the snout and lowering the hindquarters, can occur prior to exhaustion (Copp et al., 2009), it is unclear if similar gait changes occur at higher treadmill inclinations in the absence of fatigue. Only five studies utilized an incline >10° (Kemi et al., 2002; Kaurstad et al., 2012; Herbst et al., 2015; Alves et al., 2020) and no direct comparisons of treadmill incline on training responses were included in those articles. Therefore, additional research is required to confirm that treadmill inclination above 10° should be utilized for exercise training programs.

The other exercise training parameter showing subgroup differences was exercise time/session for biochemical-based outcomes. Exercise time per session varied from 30 to 90 min in the biochemical outcome group. One study with an exercise time >60 and a biochemical outcome showed a very large effect of exercise training (Vihko et al., 1979), but a wide 95% CI. Exercise times ≤30 min had a greater effect on training responses than those with training session times between 31 and 60 min. The effect size for the ≤30 min subgroup also was the largest in the combined group analysis, however, there were no significant subgroup differences between exercise times in the overall analyses. Given the relatively small number of studies per subgroup, additional research is required to confirm that shorter exercise training sessions might elicit greater biochemical training adaptations than longer individual exercise sessions.

Significant subgroup differences were observed for the exercise test used to assess changes in performance. The three most common testing protocols were increasing treadmill velocity at a fixed incline (ILT), increasing both treadmill velocity and incline at fixed intervals (GXT), and tests measuring maximal oxygen consumption ($\dot{V}O_{2max}$). Time or distance were typically used to assess performance in the ILT and GXT tests. The largest effect of exercise training was observed for $\dot{V}O_{2max}$ tests (**Figures 5**, **8**). This subgroup included six studies and had a relatively wide 95% CI (1.44–3.96). The testing protocols used to measure mouse $\dot{V}O_{2max}$ are like the protocols for ILT, increasing speed at a constant incline. But the criteria for reaching $\dot{V}O_{2max}$, e.g., a plateau in $\dot{V}O_2$ with increasing workload and respiratory exchange ratio above 1.0, is more clearly defined than those for ILT and GXT tests (e.g., time touching the shock grid or number of shocks; Poole et al., 2020). Therefore, $\dot{V}O_{2max}$ tests were placed in a separate category from ILT. The effect size for ILT was greater than that for GXT. The GXTs were primarily used by one group (Massett and Berk, 2005; Avila et al., 2017; Kim et al., 2020) and all the study protocols were 4 weeks in length. In contrast, average training duration in studies utilizing ILT protocols was 7.4 ± 2.8 weeks. Shorter duration exercise training was associated with smaller responses to exercise training (**Figures 5**, **9**) and thus, might explain some of the differences between studies utilizing ILT vs. GXT protocols. It is unclear whether differences between GXT and ILT test protocols would be observed if training programs were matched for duration. Furthermore, the combination of increasing treadmill incline and speed throughout the GXT test results in larger increases in exercise intensity at specific stages. In humans, GXT-type tests result in less uniform increases in physiological responses and more variable estimates of exercise capacity and/or oxygen consumption (Myers et al., 1991; Pescatello et al., 2014). Similar results in mice might lead to inaccurate exercise prescription and subsequently less than optimal responses to exercise training (Hoydal et al., 2007).

To further explore the contribution of moderator variables on exercise training effects, meta-regression was used to determine the role of individual variables as well as multiple variables on variation across studies. The models tested included frequency of training, treadmill velocity and incline, time per session, and training study duration to identify the exercise prescription variables most related to exercise

training outcomes. The model including all these variables accounted for 0% of between-study variance when all studies were considered. When studies were divided by outcome variables, this same model did not explain any of the between-study variance for performance-based outcomes ($R^2 = 0.0$). In contrast, this same model explained 100% of the variance for studies reporting a biochemical outcome for exercise training despite no individual variable having a $p < 0.05$. The general recommendation for meta-regression is that 10 studies should be included for each moderator variable (Baker et al., 2009). The number of studies included in the meta-regression analysis for all studies met this recommendation. However, the number of studies included with biochemical outcomes was small and therefore, the strength of the association should be interpreted with caution. Nevertheless, these results suggest that biochemical outcome variables are more strongly related to exercise training program components than are performance-based outcomes. This association implies that biochemical measurements should be incorporated into exercise training studies to provide evidence of training efficacy. This recommendation was proposed previously (Booth et al., 2010) but comes with the caveat that many of these measurements are invasive and require terminal procedures (Handschin et al., 2010).

The measurement of exercise performance in mice, including $\dot{V}O_{2max}$, is somewhat controversial. Versions of different protocols for measuring $\dot{V}O_{2max}$ in mice have been proposed in the literature, each with varying levels of evidence to support the protocol (Kemi et al., 2002; Marcaletti et al., 2011; Ayachi et al., 2016; Petrosino et al., 2016; Lemaire et al., 2017). In addition, the validity of surrogates for $\dot{V}O_{2max}$ (e.g., time to exhaustion) as estimates of exercise capacity have been questioned because of the subjective nature of the definition of volitional fatigue and/or exhaustion (Booth et al., 2010; Fuller and Thyfault, 2021) and issues with repeatability (Knab et al., 2009). Knab et al. (2009) speculated that repeatability of exercise performance measures during a maximal exercise test in mice might be related to the outcome variable and the investigator's definition of maximum. In contrast, biochemical outcomes are laboratory-based measurements with quantitative outcomes which might lead to less subjective interpretation of the outcome variable. Although some variation is likely associated with biochemical markers (Lonbro et al., 2019), standardized measurement procedures could reduce intra- and inter-investigator variation. Therefore, changes in these variables might demonstrate more consistent responses to a specific exercise intervention.

Limitations

Although data from 10 moderator variables were extracted and analyzed to explain heterogeneity between studies, there are several other factors that might influence exercise training responses. Housing temperature and time of day have been shown to influence responses to exercise and adaptations to training (Wolff and Esser, 2012; McKie et al., 2019; Sato et al., 2019). Information regarding these variables were not included as part of the data extraction process. Interest in the effect of these environmental variables on responses to exercise training is growing and subsequent analyses should consider these moderator variables. In addition, many training studies include one or more weeks of progressive increases in training load to attain a final target workload. In the current study, only the final target workload was considered for analyses. Although this early phase of the training program might influence the overall outcome, this phase was generally not well described and difficult to quantify for analytical purposes and was therefore not analyzed as part of the training program. Finally, a few studies reported subject characteristics or training paradigms as ranges. In these cases, the median value was used for any moderator variables reported as ranges to minimize missing data for any given study.

In conclusion, the results of this systematic review and meta-analysis demonstrate there is a high degree of heterogeneity across endurance exercise training studies in mice. Training duration had a significant effect of training outcome, whether the outcome was performance-based or related to biochemical traits. Parameters for exercise training prescription explained a small percentage of the variation in outcomes for performance-based traits. Therefore, investigators should consider measuring both performance and biochemical outcomes to confirm training efficacy. In addition, the lack of data on training adaptations in female mice suggests that future studies should include both male and female mice or focus solely on responses in female mice to better understand the effect of sex on exercise training responses.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

HK, CM, and MM reviewed the abstracts, titles, and full text, extracted and reviewed the data, and drafted, edited, and revised the manuscript. MM analyzed the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by funds from the Texas Tech University Office of Vice President for Research.

ACKNOWLEDGMENTS

The authors would like to thank Marc Lochbaum for CMA software access and training, and helpful discussion regarding meta-analysis procedures.

REFERENCES

Abadi, A., Crane, J. D., Ogborn, D., Hettinga, B., Akhtar, M., Stokl, A., et al. (2013). Supplementation with alpha-lipoic acid, CoQ10, and vitamin E augments running performance and mitochondrial function in female mice. *PLoS One* 8:e60722. doi: 10.1371/journal.pone.0060722

Aguiar, A. S. Jr., Tuon, T., Soares, F. S., da Rocha, L. G., Silveira, P. C., and Pinho, R. A. (2008). The effect of n-acetylcysteine and deferoxamine on exercise-induced oxidative damage in striatum and hippocampus of mice. *Neurochem. Res.* 33, 729–736. doi: 10.1007/s11064-007-9485-8

Allen, D. L., Harrison, B. C., Maass, A., Bell, M. L., Byrnes, W. C., and Leinwand, L. A. (2001). Cardiac and skeletal muscle adaptations to voluntary wheel running in the mouse. *J. Appl. Physiol.* 90, 1900–1908. doi: 10.1152/jappl.2001.90.5.1900

Almeida-Oliveira, A. R., Aquino-Junior, J., Abbasi, A., Santos-Dias, A., Oliveira-Junior, M. C., Alberca-Custodio, R. W., et al. (2019). Effects of aerobic exercise on molecular aspects of asthma: involvement of SOCS-JAK-STAT. *Exerc. Immunol. Rev.* 25, 50–62.

Alves, R. L., Cardoso, B. R. L., Ramos, I. P. R., Oliveira, B. D. S., Dos Santos, M. L., de Miranda, A. S., et al. (2019). Physical training improves exercise tolerance, cardiac function and promotes changes in neurotrophins levels in chagasic mice. *Life Sci.* 232:116629. doi: 10.1016/j.lfs.2019.116629

Alves, M. N. M., Roman-Campos, D., Parreira, A. B., Almeida, P. W. M., Cruz, J. S., Guatimosim, S., et al. (2017). Exercise training protects cardiomyocytes from deleterious effects of palmitate. *Int. J. Sports Med.* 38, 949–953. doi: 10.1055/s-0043-118010

Alves, R., Suehiro, C. L., Oliveira, F. G., Frantz, E. D. C., Medeiros, R. F., Vieira, R. P., et al. (2020). Aerobic exercise modulates cardiac NAD(P)H oxidase and the NRF2/KEAP1 pathway in a mouse model of chronic fructose consumption. *J. Appl. Physiol.* 128, 59–69. doi: 10.1152/japplphysiol.00201.2019

Avila, J. J., Kim, S. K., and Massett, M. P. (2017). Differences in exercise capacity and responses to training in 24 inbred mouse strains. *Front. Physiol.* 8:974. doi: 10.3389/fphys.2017.00974

Ayachi, M., Niel, R., Momken, I., Billat, V. L., and Mille-Hamard, L. (2016). Validation of a ramp running protocol for determination of the true VO2max in mice. *Front. Physiol.* 7:372. doi: 10.3389/fphys.2016.00372

Baker, W. L., White, C. M., Cappelleri, J. C., Kluger, J., Coleman, C. I., Health Outcomes, P., et al. (2009). Understanding heterogeneity in meta-analysis: the role of meta-regression. *Int. J. Clin. Pract.* 63, 1426–1434. doi: 10.1111/j.1742-1241.2009.02168.x

Bartalucci, A., Ferrucci, M., Fulceri, F., Lazzeri, G., Lenzi, P., Toti, L., et al. (2012). High-intensity exercise training produces morphological and biochemical changes in adrenal gland of mice. *Histol. Histopathol.* 27, 753–769. doi: 10.14670/HH-27.753

Billat, V. L., Mouisel, E., Roblot, N., and Melki, J. (2005). Inter- and intrastrain variation in mouse critical running speed. *J. Appl. Physiol.* 98, 1258–1263. doi: 10.1152/japplphysiol.00991.2004

Blair, S. N., Kohl, H. W. 3rd, Barlow, C. E., Paffenbarger, R. S. Jr., Gibbons, L. W., and Macera, C. A. (1995). Changes in physical fitness and all-cause mortality. A prospective study of healthy and unhealthy men. *JAMA* 273, 1093–1098. doi: 10.1001/jama.1995.03520380029031

Boehnke, W. H., Chernoff, G. F., and Finnell, R. H. (1987). Investigation of the teratogenic effects of exercise on pregnancy outcome in mice. *Teratog. Carcinog. Mutagen.* 7, 391–397. doi: 10.1002/tcm.1770070406

Booth, F. W., Laye, M. J., and Spangenburg, E. E. (2010). Gold standards for scientists who are conducting animal-based exercise studies. *J. Appl. Physiol.* 108, 219–221. doi: 10.1152/japplphysiol.00125.2009

Borg, M. L., Andrews, Z. B., and Watt, M. J. (2014). Exercise training does not enhance hypothalamic responsiveness to leptin or ghrelin in male mice. *J. Neuroendocrinol.* 26, 68–79. doi: 10.1111/jne.12130

Brawner, C. A., Al-Mallah, M. H., Ehrman, J. K., Qureshi, W. T., Blaha, M. J., and Keteyian, S. J. (2017). Change in maximal exercise capacity is associated with survival in men and women. *Mayo Clin. Proc.* 92, 383–390. doi: 10.1016/j.mayocp.2016.12.016

Chow, L. S., Greenlund, L. J., Asmann, Y. W., Short, K. R., McCrady, S. K., Levine, J. A., et al. (2007). Impact of endurance training on murine spontaneous activity, muscle mitochondrial DNA abundance, gene transcripts, and function. *J. Appl. Physiol.* 102, 1078–1089. doi: 10.1152/japplphysiol.00791.2006

Copp, S. W., Davis, R. T., Poole, D. C., and Musch, T. I. (2009). Reproducibility of endurance capacity and VO2peak in male Sprague-Dawley rats. *J. Appl. Physiol.* 106, 1072–1078. doi: 10.1152/japplphysiol.91566.2008

Davidson, T., Vainshelboim, B., Kokkinos, P., Myers, J., and Ross, R. (2018). Cardiorespiratory fitness versus physical activity as predictors of all-cause mortality in men. *Am. Heart J.* 196, 156–162. doi: 10.1016/j.ahj.2017.08.022

De Angelis, K., Wichi, R. B., Jesus, W. R., Moreira, E. D., Morris, M., Krieger, E. M., et al. (2004). Exercise training changes autonomic cardiovascular balance in mice. *J. Appl. Physiol.* 96, 2174–2178. doi: 10.1152/japplphysiol.00870.2003

Defina, L. F., Willis, B. L., Radford, N. B., Gao, A., Leonard, D., Haskell, W. L., et al. (2013). The association between midlife cardiorespiratory fitness levels and later-life dementia: a cohort study. *Ann. Intern. Med.* 158, 162–168. doi: 10.7326/0003-4819-158-3-201302050-00005

de Vries, R. B. M., Hooijmans, C. R., Langendam, M. W., van Luijk, J., Leenaars, M., Ritskes-Hoitinga, M., et al. (2015). A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. *Evid. Based Preclin. Med.* 2:e00007. doi: 10.1002/ebm2.7

Durigan, J. L., Peviani, S. M., Russo, T. L., Duarte, A. C., Vieira, R. P., Martins, M. A., et al. (2009a). Physical training leads to remodeling of diaphragm muscle in asthma model. *Int. J. Sports Med.* 30, 430–434. doi: 10.1055/s-0028-1112145

Durigan, J. L., Peviani, S. M., Russo, T. L., Silva, A. C., Vieira, R. P., Martins, M. A., et al. (2009b). Effects of exercise training on atrophy gene expression in skeletal muscle of mice with chronic allergic lung inflammation. *Braz. J. Med. Biol. Res.* 42, 339–345. doi: 10.1590/s0100-879x2009000400005

Egger, M., Davey Smith, G., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315, 629–634. doi: 10.1136/bmj.315.7109.629

Ferreira, J. C., Rolim, N. P., Bartholomeu, J. B., Gobatto, C. A., Kokubun, E., and Brum, P. C. (2007). Maximal lactate steady state in running mice: effect of exercise training. *Clin. Exp. Pharmacol. Physiol.* 34, 760–765. doi: 10.1111/j.1440-1681.2007.04635.x

Fiuza-Luces, C., Santos-Lozano, A., Llavero, F., Campo, R., Nogales-Gadea, G., Diez-Bermejo, J., et al. (2018). Muscle molecular adaptations to endurance exercise training are conditioned by glycogen availability: a proteomics-based analysis in the McArdle mouse model. *J. Physiol.* 596, 1035–1061. doi: 10.1113/JP275292

Foryst-Ludwig, A., Kreissl, M. C., Sprang, C., Thalke, B., Bohm, C., Benz, V., et al. (2011). Sex differences in physiological cardiac hypertrophy are associated with exercise-mediated changes in energy substrate availability. *Am. J. Physiol. Heart Circ. Physiol.* 301, H115–H122. doi: 10.1152/ajpheart.01222.2010

Fuller, K. N. Z., and Thyfault, J. P. (2021). Barriers in translating preclinical rodent exercise metabolism findings to human health. *J. Appl. Physiol.* 130, 182–192. doi: 10.1152/japplphysiol.00683.2020

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I. M., et al. (2011). American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med. Sci. Sports Exerc.* 43, 1334–1359. doi: 10.1249/MSS.0b013e318213fefb

German, E., and Hoffman-Goetz, L. (1986). The effect of cold acclimation and exercise training on cold tolerance in aged C57BL/6J mice. *J. Gerontol.* 41, 453–459. doi: 10.1093/geronj/41.4.453

Han, G. S. (2013). Endurance exercise effects on cardiac hypertrophy in mice. *J. Phys. Ther. Sci.* 25, 1525–1527. doi: 10.1589/jpts.25.1525

Handschin, C., Summermatter, S., LeBrasseur, N. K., Lin, J., and Spiegelman, B. M. (2010). For a pragmatic approach to exercise studies. *J. Appl. Physiol.* 108, 222–225. doi: 10.1152/japplphysiol.01233.2009

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., et al. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* 116, 1081–1093. doi: 10.1161/CIRCULATIONAHA.107.185649

Hellsten, Y., and Nyberg, M. (2015). Cardiovascular adaptations to exercise training. *Compr. Physiol.* 6, 1–32. doi: 10.1002/cphy.c140080

Herbst, E. A., Roussakis, C., Matravadia, S., and Holloway, G. P. (2015). Chronic treadmill running does not enhance mitochondrial oxidative capacity in the cortex or striatum. *Metabolism* 64, 1419–1425. doi: 10.1016/j.metabol.2015.07.002

Hoffman-Goetz, L., Keir, R., Thorne, R., Houston, M. E., and Young, C. (1986). Chronic exercise stress in mice depresses splenic T lymphocyte mitogenesis in vitro. *Clin. Exp. Immunol.* 66, 551–557.

Hoffman-Goetz, L., Thorne, R., Simpson, J. A., and Arumugam, Y. (1989). Exercise stress alters murine lymphocyte subset distribution in spleen, lymph nodes and thymus. *Clin. Exp. Immunol.* 76, 307–310.

Holloszy, J. O., and Coyle, E. F. (1984). Adaptations of skeletal muscle to endurance exercise and their metabolic consequences. *J. Appl. Physiol. Respir. Environ. Exerc. Physiol.* 56, 831–838. doi: 10.1152/jappl.1984.56.4.831

Hoydal, M. A., Wisloff, U., Kemi, O. J., and Ellingsen, O. (2007). Running speed and maximal oxygen uptake in rats and mice: practical implications for exercise training. *Eur. J. Cardiovasc. Prev. Rehabil.* 14, 753–760. doi: 10.1097/HJR.0b013e3281eacef1

Ingalls, C. P., Barnes, W. S., and Smith, S. B. (1996). Interaction between clenbuterol and run training: effects on exercise performance and MLC isoform content. *J. Appl. Physiol.* 80, 795–801. doi: 10.1152/jappl.1996.80.3.795

Jadeski, L., and Hoffman-Goetz, L. (1996). Exercise and in vivo natural cytotoxicity against tumour cells of varying metastatic capacity. *Clin. Exp. Metastasis* 14, 138–144. doi: 10.1007/BF00121210

Joyner, M. J., and Green, D. J. (2009). Exercise protects the cardiovascular system: effects beyond traditional risk factors. *J. Physiol.* 587, 5551–5558. doi: 10.1113/jphysiol.2009.179432

Kaurstad, G., Alves, M. N., Kemi, O. J., Rolim, N., Hoydal, M. A., Wisloff, H., et al. (2012). Chronic CaMKII inhibition blunts the cardiac contractile response to exercise training. *Eur. J. Appl. Physiol.* 112, 579–588. doi: 10.1007/s00421-011-1994-0

Kemi, O. J., Loennechen, J. P., Wisloff, U., and Ellingsen, O. (2002). Intensity-controlled treadmill running in mice: cardiac and skeletal muscle hypertrophy. *J. Appl. Physiol.* 93, 1301–1309. doi: 10.1152/japplphysiol.00231.2002

Kim, S. K., Avila, J. J., and Massett, M. P. (2020). Interaction of genetic background and exercise training intensity on endothelial function in mouse aorta. *Korean J. Physiol. Pharmacol.* 24, 53–68. doi: 10.4196/kjpp.2020.24.1.53

Kim, J., Park, J., Kim, N., Park, H. Y., and Lim, K. (2019). Inhibition of androgen receptor can decrease fat metabolism by decreasing carnitine palmitoyltransferase I levels in skeletal muscles of trained mice. *Nutr. Metab.* 16:82. doi: 10.1186/s12986-019-0406-z

Knab, A. M., Bowen, R. S., Moore-Harrison, T., Hamilton, A. T., Turner, M. J., and Lightfoot, J. T. (2009). Repeatability of exercise behaviors in mice. *Physiol. Behav.* 98, 433–440. doi: 10.1016/j.physbeh.2009.07.006

Kruger, K., Bredehoft, J., Mooren, F. C., and Rummel, C. (2016). Different effects of strength and endurance exercise training on COX-2 and mPGES expression in mouse brain are independent of peripheral inflammation. *J. Appl. Physiol.* 121, 248–254. doi: 10.1152/japplphysiol.00284.2016

Kruger, K., Gessner, D. K., Seimetz, M., Banisch, J., Ringseis, R., Eder, K., et al. (2013). Functional and muscular adaptations in an experimental model for isometric strength training in mice. *PLoS One* 8:e79069. doi: 10.1371/journal.pone.0079069

Labots, M., Laarakker, M. C., Ohl, F., and van Lith, H. A. (2016). Consomic mouse strain selection based on effect size measurement, statistical significance testing and integrated behavioral z-scoring: focus on anxiety-related behavior and locomotion. *BMC Genet.* 17:95. doi: 10.1186/s12863-016-0411-4

Lee, J. (2021). Influence of cardiorespiratory fitness on risk of dementia and dementia mortality: a systematic review and meta-analysis of prospective cohort studies. *J. Aging Phys. Act.* 29, 878–885. doi: 10.1123/japa.2019-0493

Lee, I., Huttemann, M., Kruger, A., Bollig-Fischer, A., and Malek, M. H. (2015). (−)-Epicatechin combined with 8 weeks of treadmill exercise is associated with increased angiogenic and mitochondrial signaling in mice. *Front. Pharmacol.* 6:43. doi: 10.3389/fphar.2015.00043

Lehti, T. M., Silvennoinen, M., Kivela, R., Kainulainen, H., and Komulainen, J. (2006). Effects of streptozotocin-induced diabetes and physical training on gene expression of extracellular matrix proteins in mouse skeletal muscle. *Am. J. Physiol. Endocrinol. Metab.* 290, E900–E907. doi: 10.1152/ajpendo.00444.2005

Lemaire, K. K., Thomasson, R., Noirez, P., Jaspers, R. T., and van Soest, A. J. (2017). Commentary: validation of a ramp running protocol for determination of the true VO2max in mice. *Front. Physiol.* 8:330. doi: 10.3389/fphys.2017.00330

Lightfoot, J. T., Turner, M. J., Daves, M., Vordermark, A., and Kleeberger, S. R. (2004). Genetic influence on daily wheel running activity level. *Physiol. Genomics* 19, 270–276. doi: 10.1152/physiolgenomics.00125.2004

Lightfoot, J. T., Turner, M. J., Debate, K. A., and Kleeberger, S. R. (2001). Interstrain variation in murine aerobic capacity. *Med. Sci. Sports Exerc.* 33, 2053–2057. doi: 10.1097/00005768-200112000-00012

Liu, Y. F., Chen, H. I., Yu, L., Kuo, Y. M., Wu, F. S., Chuang, J. I., et al. (2008). Upregulation of hippocampal TrkB and synaptotagmin is involved in treadmill exercise-enhanced aversive memory in mice. *Neurobiol. Learn. Mem.* 90, 81–89. doi: 10.1016/j.nlm.2008.02.005

Lonbro, S., Wiggins, J. M., Wittenborn, T., Elming, P. B., Rice, L., Pampo, C., et al. (2019). Reliability of blood lactate as a measure of exercise intensity in different strains of mice during forced treadmill running. *PLoS One* 14:e0215584. doi: 10.1371/journal.pone.0215584

Lucchetti, B. F. C., Zanluqui, N. G., de Ataides Raquel, H., Lovo-Martins, M. I., Tatakihara, V. L. H., de Oliveira Belem, M., et al. (2017). Moderate treadmill exercise training improves cardiovascular and nitrergic response and resistance to Trypanosoma cruzi infection in mice. *Front. Physiol.* 8:315. doi: 10.3389/fphys.2017.00315

Macleod, M. R., O'Collins, T., Howells, D. W., and Donnan, G. A. (2004). Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke* 35, 1203–1208. doi: 10.1161/01.STR.0000125719.25853.20

Malek, M. H., Huttemann, M., Lee, I., and Coburn, J. W. (2013). Similar skeletal muscle angiogenic and mitochondrial signalling following 8 weeks of endurance exercise in mice: discontinuous versus continuous training. *Exp. Physiol.* 98, 807–818. doi: 10.1113/expphysiol.2012.070169

Marcaletti, S., Thomas, C., and Feige, J. N. (2011). Exercise performance tests in mice. *Curr. Protoc. Mouse Biol.* 1, 141–154. doi: 10.1002/9780470942390.mo100160

Massett, M. P., and Berk, B. C. (2005). Strain-dependent differences in responses to exercise training in inbred and hybrid mice. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 288, R1006–R1013. doi: 10.1152/ajpregu.00476.2004

McKie, G. L., Medak, K. D., Knuth, C. M., Shamshoum, H., Townsend, L. K., Peppler, W. T., et al. (2019). Housing temperature affects the acute and chronic metabolic adaptations to exercise in mice. *J. Physiol.* 597, 4581–4600. doi: 10.1113/JP278221

Meeusen, R., Duclos, M., Foster, C., Fry, A., Gleeson, M., Nieman, D., et al. (2013). Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Med. Sci. Sports Exerc.* 45, 186–205. doi: 10.1249/MSS.0b013e318279a10a

Meier, P., Renga, M., Hoppeler, H., and Baum, O. (2013). The impact of antioxidant supplements and endurance exercise on genes of the carbohydrate and lipid metabolism in skeletal muscle of mice. *Cell Biochem. Funct.* 31, 51–59. doi: 10.1002/cbf.2859

Mikami, T., Sumida, S., Ishibashi, Y., and Ohta, S. (2004). Endurance exercise training inhibits activity of plasma GOT and liver caspase-3 of mice [correction of rats] exposed to stress by induction of heat shock protein 70. *J. Appl. Physiol.* 96, 1776–1781. doi: 10.1152/japplphysiol.00795.2002

Mora, S., Cook, N., Buring, J. E., Ridker, P. M., and Lee, I. M. (2007). Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. *Circulation* 116, 2110–2118. doi: 10.1161/CIRCULATIONAHA.107.729939

Murphy, S. L., Xu, J., Kochanek, K. D., Arias, E., and Tejada-Vera, B. (2021). Deaths: final data for 2018. *Natl. Vital Stat. Rep.* 69, 1–83.

Myers, J., Buchanan, N., Walsh, D., Kraemer, M., McAuley, P., Hamilton-Wessler, M., et al. (1991). Comparison of the ramp versus standard exercise protocols. *J. Am. Coll. Cardiol.* 17, 1334–1342. doi: 10.1016/s0735-1097(10)80144-5

Niebauer, J., Maxwell, A. J., Lin, P. S., Tsao, P. S., Kosek, J., Bernstein, D., et al. (1999). Impaired aerobic capacity in hypercholesterolemic mice: partial reversal by exercise training. *Am. J. Physiol.* 276, H1346–H1354. doi: 10.1152/ajpheart.1999.276.4.H1346

Niel, R., Ayachi, M., Mille-Hamard, L., Le Moyec, L., Savarin, P., Clement, M. J., et al. (2017). A new model of short acceleration-based training improves exercise performance in old mice. *Scand. J. Med. Sci. Sports* 27, 1576–1587. doi: 10.1111/sms.12809

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated

guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71

Pereira, B. C., Filho, L. A., Alves, G. F., Pauli, J. R., Ropelle, E. R., Souza, C. T., et al. (2012). A new overtraining protocol for mice based on downhill running sessions. *Clin. Exp. Pharmacol. Physiol.* 39, 793–798. doi: 10.1111/j.1440-1681.2012.05728.x

Pereira, B. C., Lucas, G., da Rocha, A. L., Pauli, J. R., Ropelle, E. R., Cintra, D., et al. (2015). Eccentric exercise leads to glial activation but not apoptosis in mice spinal cords. *Int. J. Sports Med.* 36, 378–385. doi: 10.1055/s-0034-1395589

Pereira, P. R., Oliveira-Junior, M. C., Mackenzie, B., Chiovatto, J. E., Matos, Y., Greiffo, F. R., et al. (2016). Exercise reduces lung fibrosis involving serotonin/Akt signaling. *Med. Sci. Sports Exerc.* 48, 1276–1284. doi: 10.1249/MSS.0000000000000907

Pereira, B. C., Pauli, J. R., Antunes, L. M., de Freitas, E. C., de Almeida, M. R., de Paula Venancio, V., et al. (2013). Overtraining is associated with DNA damage in blood and skeletal muscle cells of Swiss mice. *BMC Physiol.* 13:11. doi: 10.1186/1472-6793-13-11

Pereira, B. C., Pauli, J. R., De Souza, C. T., Ropelle, E. R., Cintra, D. E., Freitas, E. C., et al. (2014a). Eccentric exercise leads to performance decrease and insulin signaling impairment. *Med. Sci. Sports Exerc.* 46, 686–694. doi: 10.1249/MSS.0000000000000149

Pereira, B. C., Pauli, J. R., de Souza, C. T., Ropelle, E. R., Cintra, D. E., Rocha, E. M., et al. (2014b). Nonfunctional overreaching leads to inflammation and myostatin upregulation in Swiss mice. *Int. J. Sports Med.* 35, 139–146. doi: 10.1055/s-0033-1349077

Pescatello, L. S., Arena, R., Riebe, D., and Thompson, P. D. (eds.) (2014). *ACSM's Guidelines for Exercise Testing and Prescription*. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins.

Petrosino, J. M., Heiss, V. J., Maurya, S. K., Kalyanasundaram, A., Periasamy, M., LaFountain, R. A., et al. (2016). Graded maximal exercise testing to assess mouse cardio-metabolic phenotypes. *PLoS One* 11:e0148010. doi: 10.1371/journal.pone.0148010

Pinto, P. R., Rocco, D. D., Okuda, L. S., Machado-Lima, A., Castilho, G., da Silva, K. S., et al. (2015). Aerobic exercise training enhances the in vivo cholesterol trafficking from macrophages to the liver independently of changes in the expression of genes involved in lipid flux in macrophages and aorta. *Lipids Health Dis.* 14:109. doi: 10.1186/s12944-015-0093-3

Poole, D. C., Copp, S. W., Colburn, T. D., Craig, J. C., Allen, D. L., Sturek, M., et al. (2020). Guidelines for animal exercise and training protocols for cardiovascular studies. *Am. J. Physiol. Heart Circ. Physiol.* 318, H1100–H1138. doi: 10.1152/ajpheart.00697.2019

Robsahm, T. E., Heir, T., Sandvik, L., Prestgaard, E., Tretli, S., Erikssen, J. E., et al. (2019). Changes in midlife fitness, body mass index, and smoking influence cancer incidence and mortality: a prospective cohort study in men. *Cancer Med.* 8, 4875–4882. doi: 10.1002/cam4.2383

Rodrigues, A. C., Leal, T. F., Costa, A., Silva, F. J., Soares, L. L., Brum, P. C., et al. (2019). Effects of aerobic exercise on the inflammatory cytokine profile and expression of lipolytic and thermogenic genes in beta1-AR(−/−) mice adipose tissue. *Life Sci.* 221, 224–232. doi: 10.1016/j.lfs.2019.02.031

Sato, S., Basse, A. L., Schönke, M., Chen, S., Samad, M., Altıntaş, A., et al. (2019). Time of exercise specifies the impact on muscle metabolic pathways and systemic energy homeostasis. *Cell Metab.* 30, 92.e4–110.e4. doi: 10.1016/j.cmet.2019.03.013

Savage, K. J., and McPherron, A. C. (2010). Endurance exercise training in myostatin null mice. *Muscle Nerve* 42, 355–362. doi: 10.1002/mus.21688

Sousa, A. S., Sponton, A. C. S., Trifone, C. B., and Delbin, M. A. (2019). Aerobic exercise training prevents perivascular adipose tissue-induced endothelial dysfunction in thoracic aorta of obese mice. *Front. Physiol.* 10:1009. doi: 10.3389/fphys.2019.01009

Steiner, J. L., Murphy, E. A., McClellan, J. L., Carmichael, M. D., and Davis, J. M. (2011). Exercise training increases mitochondrial biogenesis in the brain. *J. Appl. Physiol.* 111, 1066–1071. doi: 10.1152/japplphysiol.00343.2011

Sturgeon, K., Muthukumaran, G., Ding, D., Bajulaiye, A., Ferrari, V., and Libonati, J. R. (2015). Moderate-intensity treadmill exercise training decreases murine cardiomyocyte cross-sectional area. *Physiol. Rep.* 3:e12406. doi: 10.14814/phy2.12406

Sui, X., Sarzynski, M. A., Lee, D. C., Lavie, C. J., Zhang, J., Kokkinos, P. F., et al. (2017). Longitudinal patterns of cardiorespiratory fitness predict the development of hypertension among men and women. *Am. J. Med.* 130, 469.e2–476.e2. doi: 10.1016/j.amjmed.2016.11.017

Suominen, H., Kiiskinen, A., and Heikkinen, E. (1980). Effects of physical training on metabolism of connective tissues in young mice. *Acta Physiol. Scand.* 108, 17–22. doi: 10.1111/j.1748-1716.1980.tb06495.x

Toti, L., Bartalucci, A., Ferrucci, M., Fulceri, F., Lazzeri, G., Lenzi, P., et al. (2013). High-intensity exercise training induces morphological and biochemical changes in skeletal muscles. *Biol. Sport* 30, 301–309. doi: 10.5604/20831862.1077557

Troxell, M. L., Britton, S. L., and Koch, L. G. (2003). Selected contribution: variation and heritability for the adaptational response to exercise in genetically heterogeneous rats. *J. Appl. Physiol.* 94, 1674–1681. doi: 10.1152/japplphysiol.00851.2002

Uddin, G. M., Youngson, N. A., Sinclair, D. A., and Morris, M. J. (2016). Head to head comparison of short-term treatment with the NAD(+) precursor Nicotinamide mononucleotide (NMN) and 6 weeks of exercise in obese female mice. *Front. Pharmacol.* 7:258. doi: 10.3389/fphar.2016.00258

Vesterinen, H. M., Sena, E. S., Egan, K. J., Hirst, T. C., Churolov, L., Currie, G. L., et al. (2014). Meta-analysis of data from animal studies: a practical guide. *J. Neurosci. Methods* 221, 92–102. doi: 10.1016/j.jneumeth.2013.09.010

Vieira, R. P., de Andrade, V. F., Duarte, A. C., Dos Santos, A. B., Mauad, T., Martins, M. A., et al. (2008). Aerobic conditioning and allergic pulmonary inflammation in mice. II. Effects on lung vascular and parenchymal inflammation and remodeling. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 295, L670–L679. doi: 10.1152/ajplung.00465.2007

Vihko, V., Salminen, A., and Rantamaki, J. (1979). Exhaustive exercise, endurance training, and acid hydrolase activity in skeletal muscle. *J. Appl. Physiol. Respir. Environ. Exerc. Physiol.* 47, 43–50. doi: 10.1152/jappl.1979.47.1.43

Waters, R. E., Rotevatn, S., Li, P., Annex, B. H., and Yan, Z. (2004). Voluntary running induces fiber type-specific angiogenesis in mouse skeletal muscle. *Am. J. Physiol. Cell Physiol.* 287, C1342–C1348. doi: 10.1152/ajpcell.00247.2004

Wernig, A., Salvini, T. F., and Irintchev, A. (1991). Axonal sprouting and changes in fibre types after running-induced muscle damage. *J. Neurocytol.* 20, 903–913. doi: 10.1007/BF01190468

Wolff, G., and Esser, K. A. (2012). Scheduled exercise phase shifts the circadian clock in skeletal muscle. *Med. Sci. Sports Exerc.* 44, 1663–1670. doi: 10.1249/MSS.0b013e318255cf4c

Woods, J. A., Ceddia, M. A., Zack, M. D., Lowder, T. W., and Lu, Q. (2003). Exercise training increases the naive to memory T cell ratio in old mice. *Brain Behav. Immun.* 17, 384–392. doi: 10.1016/S0889-1591(03)00030-8

Zaccardi, F., O'Donovan, G., Webb, D. R., Yates, T., Kurl, S., Khunti, K., et al. (2015). Cardiorespiratory fitness and risk of type 2 diabetes mellitus: a 23-year cohort study and a meta-analysis of prospective studies. *Atherosclerosis* 243, 131–137. doi: 10.1016/j.atherosclerosis.2015.09.016

Zhang, P., Sui, X., Hand, G. A., Hebert, J. R., and Blair, S. N. (2014). Association of changes in fitness and body composition with cancer mortality in men. *Med. Sci. Sports Exerc.* 46, 1366–1374. doi: 10.1249/MSS.0000000000000225

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Massett, Matejka and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.