

Supplementary Material

The effect of exercise training on intrahepatic triglyceride and hepatic insulin sensitivity: a systematic review and meta-analysis

Jack A. Sargeant^{1,2}, Laura J. Gray³, Danielle H. Bodicoat^{2,4}, Scott A. Willis^{1,2}, David J. Stensel^{1,2}, Myra A. Nimmo⁵, Guruprasad P. Aithal^{6,7}, James A. King^{1,2}.

¹National Centre for Sport and Exercise Medicine, School of Sport, Exercise & Health Sciences, Loughborough University, UK

²National Institute for Health Research (NIHR) Leicester Biomedical Research Centre, Leicester, UK

³Department of Health Sciences, University of Leicester, UK

⁴Diabetes Research Centre, University of Leicester, UK

⁵College of Life & Environmental Sciences, University of Birmingham, UK

⁶Nottingham Digestive Diseases Centre, School of Medicine, University of Nottingham

⁷National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, UK

Address for correspondence:

Dr James King,

Lecturer in Exercise Physiology

School of Sport, Exercise and Health Sciences, Loughborough University,

United Kingdom, LE11 3TU

Phone: +44(0)1509 228457

Email: j.a.king@lboro.ac.uk

Supplementary Methods

Literature Search Terms

The following search terms were used. All terms *within* brackets were combined using ‘OR’:

(exercise, physical activity, training) AND [(liver fat, hepatic fat, intrahepatic triglyceride, IHTG, intrahepatocellular lipids, intrahepatic lipids, non-alcoholic fatty liver disease, NAFLD, non-alcoholic fatty liver, fatty liver, hepatic steatosis, non-alcoholic steatosis, liver steatosis, non-alcoholic steatohepatitis, NASH, hepatic steatohepatitis, liver steatohepatitis, liver function) OR (hepatic insulin sensitivity, hepatic insulin resistance, liver insulin sensitivity, liver insulin resistance, hepatic IR, liver IR, hepatic glucose production, liver glucose production, endogenous glucose production, glycolysis, gluconeogenesis)].

Data extraction

Imputation of the mean and standard deviation (SD) during data extraction

When the variation of change from pre- to post-intervention was reported as standard error or confidence intervals, established statistical equations were used to convert these to SD. When only pre- and post-intervention data were available, the mean change and SD were imputed as previously reported (1), assuming a correlation coefficient for change in IHTG of 0.80 (2–4).

Data presented in graphical form only

When data were presented in graphical form only, values were estimated using commercially available software (Digitizeit, Version 2.2, Bormann, I., Braunschweig, Germany).

Data presented as median and interquartile range

When original data were reported as median and interquartile range, the median change was extracted and used in place of the mean, whilst the IQR was divided by 1.35 as an estimate SD (1). A sensitivity analysis removing these studies suggested that using data in this manner had no substantial impact on the results reported.

Converting data to a consistent unit of measurement

Data presented in this meta-analysis are, as much as possible, presented in a consistent format using similar units of measurement. When IHTG was reported as the ratio between intracellular lipid and water (lipid:water), this was converted to liver fat fraction using the following calculation:

$$(\text{lipid:water}) / [1+(\text{lipid:water})] \times 100$$

In some cases, data were converted during extraction using conversion factors as follows:

$$\text{Insulin: } 1 \text{ mU} = 6 \text{ pmol}$$

$$\text{Glucose: } 1 \text{ mmol}\cdot\text{L}^{-1} = 18 \text{ mg}\cdot\text{dL}^{-1}$$

$$\text{EGP: } 1 \text{ }\mu\text{mol}\cdot\text{kg}^{-1}\cdot\text{min}^{-1} = 0.18 \text{ mg}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$$

When data were also normalised, such as when presenting aerobic capacity in $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ rather than $\text{L}\cdot\text{min}^{-1}$, the mean value for the whole study group at baseline was used.

Inclusion of subsets of participants

When a paper reported one of the primary outcomes of this review in a subset of individuals (or when only a subset of individuals met the inclusion criteria), only these individuals were included. If descriptive data specific to this subset were available, they were extracted. If not,

the sample size was adjusted to represent the subset of individuals but the mean descriptive data for the whole study population was used. This applies to the following studies:

Cuthbertson et al., 2016 (5): only 19 individuals (ex: 12; con: 7) underwent hyperinsulinaemic, euglycaemic clamp and were thus included in HISI and percentage suppression of EGP analyses.

Hickman et al., 2013 (6): only nine of the 13 patients in the exercise group underwent post-training liver biopsy and were thus included in IHTG analyses. Furthermore, complete data to allow calculation of HISI were only available for 11 patients.

Keating et al., 2015 (7): three separate exercise groups completing different exercise interventions were included in this study. One of these groups was ineligible as the mean baseline IHTG was $< 5.56\%$.

Langley et al., 2016 (8): only the dysglycaemic individuals had IHTG $> 5.56\%$ at baseline.

Meex et al., 2010 (9): Raw data was collected for the calculation of HISI. One participant was removed from analysis because no basal EGP data was available.

Sargeant et al., 2018 (10): only eight of the nine participants completed hyperinsulinaemic, euglycaemic clamps and so only these participants were included in HISI and percentage suppression of EGP analyses.

van der Heijden et al., 2010 (11): only seven individuals had NAFLD (defined as liver fat $> 5.6\%$) and so only this subset was included in the hepatic steatosis analysis. All participants were included in the HISI analysis as NAFLD diagnosis was not an inclusion criterion for this outcome.

Inclusion of studies with multiple follow-up measurements

One manuscript (12) reported data after six months and 12 months of intervention in two separate exercise groups. The interventions differed only during the first six months and, therefore, these data were extracted and used.

Risk of bias assessment

The Downs and Black scale was modified in two ways. Blinding participants and experimenters to group allocations is a difficult task in exercise trials. It is possible, however, to blind experimenters who are conducting data analysis. As such, item 14 of the original scale (concerning participant blinding) was removed, whilst item 15 (concerning experimenter blinding) was scored according to whether attempts were made to blind experimenters during data analysis. Item 27, concerning statistical power, was modified as follows:

Formal power calculation performed based on detecting a significant change in IHTG = score of two awarded

Formal power calculation performed based on detecting a significant difference in a relevant and related outcome but which was not IHTG = score of one awarded

No formal power calculation performed = score of zero awarded.

Keating *et al* (7) performed a power calculation on IHTG to detect change *between* groups indicating that 116 participants were required per group. However, due to limited data to inform this power calculation and the difficulties in performing a study of that design in such large numbers they did not recruit to this extent. Given that 1) the purpose of this meta-analysis was not to determine differences *between* groups, 2) they had attempted to perform

an appropriate calculation and 3) the numbers they did recruit were similar to those of other suitably powered studies, a score of two was awarded for item 27 (statistical power).

Combination of groups

In RCTs with multiple intervention groups, the intervention groups were combined, where possible, using appropriate statistical formulae (1) and, for the purpose of exercise programme description, the exercise intensity, duration and frequency of the combined group were calculated as the weighted mean of the individual groups.

Supplementary Results

Table S1 – Participant characteristics and outcome measures of studies assessing changes in IHTG

Ref. <i>Study Design</i>	Intervention Overview	Participant Characteristics at Baseline							Relative Change in Body Weight (%)	Comments
		N (M/F)	Age (years)	BMI (kg•m ⁻²)	Body Fat (%)	Aerobic Capacity (ml•kg ⁻¹ •min ⁻¹)	Baseline IHTG (%)	Clinical Conditions		
Cassidy et al. (2016) <i>RCT</i>	HIIT 3 times per week for 12 weeks.	Ex: 12 (10/2) Con: 11 (8/3)	61 ± 9 59 ± 9	31 ± 5 32 ± 6	35.4 (est.) 39.6 (est.)	21.8 ± 5.4 20.3 ± 6.1	6.9 ± 6.9 7.1 ± 6.8	Hepatic steatosis according to mean baseline measures. Diagnosed T2DM with stable control for ≥ 6 months.	-1.1 [#] 1.1	<i>Reported medications:</i> Metformin (Ex: n=7, Con: n=7) Statins (Ex: n=7, Con: n=6) Antihypertensives (Ex: n=3, Con: n=5)
Cuthbertson et al. (2016) <i>RCT</i>	Aerobic 3 – 5 times per week for 12 weeks.	Ex: 30 (23/7) Con: 20 (16/4)	50 (46–58) 52 (46–59)	30.6 (29.0–32.9) 29.7 (28.0–33.0)	30.4 (25.9–32.1) 31.0 (26.5–37.7)	23.7 (21.7–27.8) 32.2 (20.9–25.6)	19.4 (14.6–36.1) 16.0 (9.6–32.5)	Diagnosed NAFLD	-2.5 (-3.5 – -1.4) ^{a,#} 0.2 (-0.8–1.1)	<i>Notable exclusion criteria:</i> T2DM and IHD Excessive weight loss during the course of the study
Hallsworth et al. (2011) <i>RCT</i>	Resistance 3 times per weeks for 8 weeks.	Ex: 11 (NR) Con: 8 (NR)	52 ± 13 62 ± 7	32.3 ± 4.9 32.3 ± 4.8	37 ± 8 41 ± 6	21.8 ± 3.8 18.5 ± 5.2	14.0 ± 9.1 11.2 ± 8.4	Clinically defined but non-advanced NAFLD (defined as IHTG > 5% with NAFLD fibrosis score < -1.445). T2DM accepted provided diet and metformin prescription were stable for ≥ 6 months.	0.0 0.6	<i>Notable exclusion criteria:</i> Exogenous insulin therapy. IHD Weight loss > 2.5% during the study. <i>Other:</i> Exercise group were significantly younger than control (<i>p</i> < 0.05).
Hallsworth et al. (2015) <i>RCT</i>	HIIT 3 times per week for 12 weeks.	Ex: 12 (6/6) Con: 11 (10/1)	54 ± 10 52 ± 12	31.0 ± 4.0 31.0 ± 5.0	38.4 ± 6.4 34.5 ± 7.0	21.9 ± 6.2 24.6 ± 5.7	10.6 ± 4.9 10.3 ± 4.4	Clinically defined but non-advanced NAFLD (defined as IHTG > 5% with NAFLD fibrosis score < -1.445). Metformin accepted but participants with any other medication for T2DM were excluded.	-1.6 ^{a,#} 0.0	None
Haus et al. (2013) <i>Uncontrolled Intervention</i>	Aerobic Daily for 1 week.	Ex: 17 (NR)	54 ± 2	34.4 ± 1.0	NR	24.3 ± 1.5	19.4 ± 3.3	Diagnosed NAFLD.	0.2	<i>Notable exclusion criteria:</i> T1DM, T2DM and CVD

Hickman et al. (2013) <i>RCT*</i>	Resistance 3 times per week for 24 weeks.	Ex: 9 (7/2)	48 ± 9	33.6 ± 5.8	38.8 ± 7.6	21.6 ± 7.3	71 ± 32 [†]	Diagnosed NAFLD. 85% met diagnostic criteria for NASH.	-2.6 ± 4.65	<i>Notable exclusion criteria:</i> Diabetes Alcohol consumption > 40 and 20 g•day ⁻¹ for men and women respectively.
Houghton et al. (2017) <i>RCT</i>	Combined (HIIT + Resistance) 3 times per week for 12 weeks.	Ex: 12 (7/5) Con: 12 (7/5)	54 ± 12 51 ± 16	33.0 ± 7.0 33.0 ± 5.0	NR	25.0 ± 8.0 21.0 ± 5.0	12.0 ± 9.0 10.0 ± 5.0	Histologically characterised NASH with NAFLD activity score ≥ 5.	1.1 1.1	<i>Notable exclusion criteria:</i> ≥60 minutes moderate-vigorous physical activity per week Insulin sensitising treatments Cardiac or renal diseases.
Johnson et al. (2009) <i>RCT</i>	Aerobic 3 times per week for 4 weeks.	Ex: 12 (NR) Con: 7 (NR)	49 ± 8 47 ± 10	32.2 ± 2.8 31.1 ± 2.9	NR	25.9 ± 4.8 25.0 ± 4.2	8.6 ± 8.6 9.2 ± 10.1	Hepatic steatosis according to mean baseline measures. Hypertension allowed.	-0.3 -0.2	<i>Notable exclusion criteria:</i> Lipid lowering medications FPG ≥ 7.0 mmol•L ⁻¹ <i>Reported medications:</i> Anti-hypertensive medications (n=5) Note – medications were unaltered for the duration of the study except when participants refrained for 72 hours prior to study assessments.
Keating et al. (2015) <i>RCT</i>	Aerobic 3-4 times per week for 8 weeks.	Ex1(HI:LO): 12 (6/6) Ex2(LO:HI): 11 (5/6) Con: 12 (3/9)	44 ± 10 45 ± 9 39 ± 10	36.3 ± 5.9 34.0 ± 3.1 32.2 ± 4.8	NR	21.9 ± 4.8 24.5 ± 3.0 21.7 ± 6.2	8.4 ± 5.2 9.4 ± 6.6 7.7 ± 9.0	Hepatic steatosis according to mean baseline measures.	-1.2 ^{a,#} -1.5 ^{a,#} 0.9 ^a	<i>Notable exclusion criteria:</i> Reported exercise on >3 days per week. Lipid-lowering or insulin-sensitizing medications.
Langley et al. (2016) <i>Uncontrolled Intervention</i>	Combined (Aerobic + HIIT + Resistance) 4 times per week for 12 weeks	Ex: 11 (11/0)	53 (10)	27.8 (5.3)	NR	38.7 (8.1)	11.0 (11.9)	Hepatic steatosis according to mean baseline measures Dysglycaemic according to OGTT performed during screening.	-1.2	<i>Notable exclusion criteria:</i> Structured exercise performed > once per week Hypertension, other liver or kidney diseases, chronic inflammatory disease or medications known to affect glucose metabolism.
Malin et al. (2013) <i>Uncontrolled Intervention</i>	Aerobic Daily for 1 week.	Ex: 13 (6/7)	51 ± 12	33.3 ± 3.2	NR	24.9 ± 5.4	23.1 ± 14.8	Hepatic steatosis confirmed during baseline measures (defined as > 5% IHTG).	0.6	None

Oh et al. (2014) <i>Uncontrolled Intervention</i>	Acceleration / Vibration 3 times per week for 6 weeks.	Ex: 18 (4/14)	NR	28.2 (25.5–33.2)	NR	NR	12.2 (5.4–20.6)	Diagnosed NAFLD by medical history, serum ALT and ultrasound. Confirmed by baseline IHTG > 5%.	-0.4 ^a	None
Pugh et al. (2014) <i>RCT</i>	Aerobic 3 – 5 times per week for 16 weeks.	Ex: 13 (7/6) Con: 8 (4/4)	50 ± 3 47 ± 5	30.0 ± 0.8 30.0 ± 2.0	NR	26.4 ± 2.3 27.0 ± 2.8	21.3 ± 12.8 19.2 ± 6.1	Clinically diagnosed NAFLD defined as IHTG ≥ 5.5% Individuals taking anti-hypertensive medications were allowed.	-2.4 ± 2.0 -1.1 ± 2.0	<i>Notable exclusion criteria:</i> Any form of structured exercise or > 2 hours of low-intensity physical activity per week. T2DM, IHD, habitual smokers. <i>Reported medications:</i> Anti-hypertensive medications (Ex: n=9) Note - medications were unaltered throughout the duration of the study.
Sargeant et al. (2018) <i>Controlled Longitudinal Intervention</i>	HIIT 3 times per week for 6 weeks.	Ex: 9 (9/0)	41 ± 8	31.7 ± 3.1	28.7 ± 3.0	31.8 ± 4.8	15.6 ± 8.3	NAFLD defined as IHTG ≥ 5.56 percent in the absence of reported secondary causes as determined during participant screening.	-1.2	Participants were weight stable <i>Notable exclusion criteria:</i> Any form of diagnosed metabolic disease or taking medication known to influence lipid metabolism or glycaemic control.
Sullivan et al. (2012) <i>RCT</i>	Aerobic 5 times per week for 16 weeks.	Ex: 12 (4/8) Con: 6 (1/5)	49 ± 8 48 ± 8	37.1 ± 3.8 40.0 ± 5.4	38.9 ± 2.1 42.5 ± 3.6	22.8 ± 4.5 18.5 ± 7.1	20.2 ± 14.6 21.4 ± 21.6	NAFLD defined as IHTG > 10%.	-0.2 ^a 0.2 ^a	<i>Notable exclusion criteria:</i> T2DM or plasma TG > 400mg•dL ⁻¹ . Not weight-stable (> 3kg weight change in previous three months). Self-reported exercise > 1 hour per week.
van der Heijden et al. (2010a) [‡] <i>Uncontrolled Intervention</i>	Aerobic 4 times per week for 12 weeks.	Ex: 15 (7/8)	16 ± 2	33.7 ± 4.3	38.3 ± 5.8	26.8 ± 6.3	9.0 ± 12.0	Hepatic steatosis according to mean baseline data.	-0.5	All participants were post-pubertal. Participants were obese for ≥ 5 years and weight stable for ≥ 6 months. <i>Notable exclusion criteria:</i> Participation in organised school athletic programme or ≥ 45 minutes self-reported light to moderate physical activity. Any form of metabolic disease. Any medication (including contraceptives). 1 st degree relatives with diabetes. Morbid obesity (>50% body fat).

van der Heijden et al. (2010b) <i>Uncontrolled Intervention</i>	Resistance 2 times per week for 12 weeks.	Ex: 7 (NR)	16 ± 1	35.3 ± 1.9	42.6 ± 5.3	NR	13.9 ± 11.4	Hepatic steatosis according to mean baseline measures but no formal diagnosis of NAFLD reported.	2.6 ^a	All participants were post-pubertal. Participants were obese for ≥5 years and weight stable for ≥6 months. <i>Notable exclusion criteria:</i> Participation in any organised school athletic programme or ≥ 45 minutes self-reported light to moderate physical activity. Any form of metabolic disease. Any medication (including contraceptives). 1 st degree relatives with diabetes. Morbid obesity (>50% body fat).
Zhang et al (2016) <i>RCT</i>	Aerobic 5 times per week for 24 weeks.	Ex1(Mod): 73 (22/51) Ex2(Vig): 73 (21/52) Con: 74 (28/46)	54 ± 7 53 ± 7 54 ± 7	28.1 ± 3.3 27.9 ± 2.7 28.0 ± 2.7	33.5 ± 5.5 34.8 ± 5.3 33.7 ± 7.1	NR	18.0 ± 9.9 18.4 ± 9.9 17.5 ± 11.0	NAFLD diagnosed initially by ultrasound and the confirmed by ¹ H-MRS during screening	-2.8 ^{a,#} -6.0 ^{a,#} -2.1 ^a	<i>Notable exclusion criteria:</i> History of other chronic liver diseases, hypertension, chronic kidney disease, hyperthyroidism, myocardial infarction (within 6 month) or heart failure. Participation in weight loss programmes

Data presented as mean ± SD or median (IQR); Sample sizes represent the number of individuals entered into analyses; *Study did not include ‘standard care’ or no ‘intervention group’.

Patients were randomised to either exercise or dietary interventions; † Liver biopsy used. This number represent the percentage of hepatocytes affected by steatosis; ‡ Manuscript refers to same study as van der Heijden et al (2009); ^a significant difference from baseline ($P < 0.05$); [#] significant interaction between exercise and control groups ($P < 0.05$); ¹H-MRS: Magnetic resonance spectroscopy; ALT: Alanine aminotransferase; BMI: Body mass index; Con: control group only; CVD: Cardiovascular disease; Ex: Exercise group only; FPG: fasting plasma glucose; HIIT: High-intensity interval training; IHD: Ischaemic heart disease; IHTG: Intrahepatic triglyceride; NAFLD: Non-alcoholic fatty liver disease; NASH: Non-alcoholic steatohepatitis; NR: Not reported; RCT: Randomised controlled trial; T1DM: Type 1 diabetes mellitus; T2DM: Type 2 diabetes mellitus; TAG: Triacylglycerol; Mod: Moderate; Vig: Vigorous.

Table S2 - Participant characteristics and outcome measures of studies assessing changes in HISI and %EGP_{supp}.

Ref. <i>Study Design</i>	Intervention Overview	Participant Characteristics at Baseline										Comments
		N (M:F)	Age (years)	BMI (kg•m ⁻²)	Body Fat (%)	Aerobic Capacity (ml•kg ⁻¹ •min ⁻¹)	Fasted Glucose (mmol•L ⁻¹)	Fasted Insulin (μU•mL ⁻¹)	HISI (mg•m ⁻² •min ⁻¹ per μU•mL ⁻¹)	EGP suppression (%)	Clinical Conditions	
Cuthbertson et al. (2016) <i>RCT</i>	Aerobic 3 – 5 times per week for 12 weeks.	Ex: 12 (8/3/1NR) Con: 7 (3/4)	44 ± 13 50 ± 12	31.0 ± 2.3 29.0 ± 3.2	30.4 (25.9-32.1) 31.0 (26.5-37.7)	28.1 ± 7.2 24.2 ± 11.3	4.79 ± 0.42 4.84 ± 0.64	17.4 ± 11.1 13.4 ± 4.9	1.11 ± 0.60 1.00 ± 0.58	50.1 ± 20.2 46.5 ± 27.3	Diagnosed NAFLD.	<i>See Table 4.2 for Notable exclusion criteria</i> <i>Insulin Dose and Infusion Duration:</i> 0.3 mU•kg ⁻¹ •min ⁻¹ for 120 minutes. <i>Potential Confounding Variables:</i> Significant reduction in body weight and IHTG in the exercise group.
Hickman et al. (2013) <i>RCT*</i>	Resistance 3 times per week for 24 weeks.	Ex: 13 (9/4)	50 ± 9	33 ± 6	39 ± 8	NR	5.5 ± 0.5	24 ± 22	0.85 ± 0.61	Not Measured	Diagnosed NAFLD. 85% met diagnostic criteria for NASH	<i>See Table 4.2 for Notable exclusion criteria</i> NOTE: %EGP _{supp} was reported at high-dose insulin infusion (1 mU•kg ⁻¹ •min ⁻¹) and was improved by exercise training.
Lee et al. (2013) <i>RCT</i>	Ex 1: Aerobic Ex 2: Resistance 3-4 times per week for 8 weeks.	Ex1: 16 (0/16) Ex2: 16 (0/16) Con: 12 (0/12)	15 ± 2 15 ± 2 15 ± 2	32.9 ± 3.8 36.4 ± 3.8 35.3 ± 4.0	47.8 ± 4.2 51.5 ± 4.7 51.3 ± 3.5	28.5 ± 3.8 24.3 ± 4.3 23.9 ± 3.0	5.18 ± 0.33 5.24 ± 0.38 5.39 ± 0.32	28.6 ± 16.5 45.8 ± 22.0 31.1 ± 15.3	24.2 ± 12.2 [†] 16.5 ± 10.5 [†] 22.9 ± 14.4 [†]	Not Measured	None reported.	<i>Notable exclusion criteria:</i> Endocrine disorders (PCOS, T2DM) Medication known to influence glucose metabolism or body composition. <i>Potential Confounding Variables:</i> No change in body weight in either group. IHTG significantly reduced in aerobic exercise group only.

Meex et al. (2010) <i>Uncontrolled Intervention</i>	Combined (Aerobic + Resistance) 3 times per week (2 Aerobic, 1 Resistance) for 12 weeks.	Ex1: 20 (20/0) Ex2: 17 (17/0)	59 ± 1 59 ± 1	29.7 ± 3.6 30.0 ± 3.4	31.5 ± 1.4 31.1 ± 1.4	28.8 ± 4.5 27.5 ± 5.1	5.90 ± 0.45 9.00 ± 1.70	18.1 ± 10.7 17.0 ± 5.3	1.1 ± 0.5 0.8 ± 0.3	Not Measured	<p><i>Notable exclusion criteria:</i> Cardiac disease, impaired liver or renal function, BMI > 35kg•m⁻². Exogenous insulin therapies.</p> <p><i>Medications:</i> All T2DM participants were on oral antidiabetic agents. Medication was unchanged throughout the duration of the study but discontinued for 7 days prior to each clamp assessment.</p> <p><i>Potential Confounding Variables:</i> No change in body weight in either group. IHTG not measured.</p>
Sargeant et al. (2018) <i>Controlled Longitudinal Intervention</i>	HIIT 3 times per week for 6 weeks.	Ex: 8 (8/0)	40 ± 7	31.5 ± 3.3	28.1 ± 2.5	32.0 ± 5.1	4.7 ± 0.3	17.6 ± 4.5	0.68 ± 0.20	59.9 ± 17.4	<p>NAFLD defined as IHTG ≥ 5.56 percent in the absence of reported secondary causes as determined during participant screening.</p> <p><i>See Table 4.2 for Notable exclusion criteria</i></p> <p><i>Insulin Dose and Infusion Duration:</i> 20 mU•m⁻²•min⁻¹ for 120 minutes.</p> <p><i>Potential Confounding Variables:</i> No change in body weight. IHTG significantly reduced from pre- to post-training.</p>

Shojaee-Moradie et al. (2007) <i>RCT</i>	Aerobic 3 times per week for 6 weeks.	Ex: 10 (10/0) Con: 7 (7/0)	47 ± 9 55 ± 11	27.6 ± 1.9 27.6 ± 2.4	25.6 (est.) 24.4 (est.)	31.0 ± 3.2 27.0 ± 5.3	NR NR	10.6 ± 6.3 11.4 ± 7.6	2.09 ± 1.03 1.91 ± 0.87	57.0 ± 15.1 62.0 ± 18.9	None reported.	<i>Notable exclusion criteria:</i> T2DM, hyperlipidaemia Lipid-lowering medications Those already engaged in regular physical activity. <i>Insulin Dose and Infusion Duration:</i> 0.3 mU•kg ⁻¹ •min ⁻¹ for 120 minutes. <i>Potential Confounding Variables:</i> No change in either body weight or IHTG in either group.
van der Heijden et al. (2009) [‡] <i>Uncontrolled Intervention</i>	Aerobic 4 times per week for 12 weeks.	Ex: 15 (7/8)	16 ± 2	33.2 ± 3.5	38.3 ± 5.8	27.5 ± 6.3	5.0 ± 0.4	20.2 ± 9.6	0.87 ± 0.43	Not Measured	Hepatic steatosis according to mean baseline measures.	<i>See Table 4.2 for Notable exclusion criteria</i> <i>Potential Confounding Variables:</i> No change in body weight. IHTG not measured.
van der Heijden et al. (2010b) <i>Uncontrolled Intervention</i>	Resistance 2 times per week for 12 weeks.	Ex: 12 (6/6)	16 ± 2	35.3 ± 2.4	42.6 ± 5.3	NR	5.10 ± 0.35	23.0 ± 6.4	0.63 ± 0.18	Not Measured	Hepatic steatosis according to mean baseline measures.	<i>See Table 4.2 for Notable exclusion criteria</i> <i>Standardisation:</i> Post-intervention assessments were performed 3 days after the final exercise training session. <i>Potential Confounding Variables:</i> Body weight significantly increased from baseline. Much of this was accounted for by an increase in LBM. IHTG not measured.

Data presented as mean ± SD or median (IQR); Samples sizes represent the number of individuals entered into analyses; *Study did not include 'standard care' or no 'intervention group'.

Patients were randomised to either exercise or dietary interventions; † HISI presented as mg•kg⁻¹•min⁻¹ per μU•mL⁻¹; ‡ Manuscript refers to same study as van der Heijden et al (2010a); #

baseline values not reported and unavailable upon request; BMI: Body mass index; Con: control group only; EGP: Endogenous glucose production; Ex: Exercise group only; HISI:

Hepatic insulin sensitivity index; IHTG: Intrahepatic triglyceride; NAFLD: Non-alcoholic fatty liver disease; PCOS: Polycystic ovary syndrome; RCT: Randomised controlled trial; T2DM: Type 2 diabetes mellitus.

Table S3 – Intervention characteristics of all included studies

Ref. <i>Study Design</i>	Exercise Type	Intervention Duration	Session Frequency	Details of Exercise intervention	Exercise Supervision	Instructions to Control Groups / Details of Placebo Intervention	Other Instructions
Cassidy et al. (2016) <i>RCT</i>	HIIT	12 weeks	3 times per week	Sessions consisted of 5 cycling intervals at an intensity equating to 'very hard' (16-17) on a Borg RPE scale. Interval length progressed by 10 seconds per week from 2 minutes in week 1 to 3 minutes 50 seconds by week 12. Intervals were interspersed with 3 minutes consisting of 90 seconds passive recovery, 60 seconds upper body resistance band exercise and 30 seconds preparation for the subsequent interval.	The initial exercise session was supervised by a member of the research team after which instructions were provided via voice-recordings loaded onto an iPod (Apple, CA, USA). Adherence was monitored via exercise diaries.	Continued standard care.	Participants were instructed to continue their normal routine and medical care, making no changes to diet, habitual activity or medication.
Cuthbertson et al. (2016) <i>RCT</i>	Aerobic	12 weeks	3 – 5 times per week	Participants were given the choice of exercising on a treadmill, cross-trainer, cycle ergometer or rower. Intensity increased from 30% HRR at week 1 to 60% HRR by week 12. Frequency and duration progressed from 30 minutes, 3 times per week during week 1 to 45 minutes, 5 times per week by week 12.	One session per week was supervised by a trained exercise physiologist. The remaining sessions monitored via Wellness System™ (Technogym U.K. Ltd.) or by repeated telephone/email contact.	Education and advice about the health benefits of exercise in NAFLD.	Participants in the exercise groups were instructed to make no dietary modifications (confirmed by 3-day self-report food diaries) To avoid disturbance to behaviour, participants in the control group were given no instructions regarding diet or lifestyle.
Hallsworth et al. (2011) <i>RCT</i>	Resistance	8 weeks	3 times per week	Sessions consisted of 8 whole-body exercises targeting large muscle groups performed as a circuit. Sessions progressed from 2 circuits at 50% 1RM during week 1 to 3 circuits at 70% 1RM by week 7. Participants were encouraged to increase the resistance each week if possible.	Sessions were supervised biweekly. Heart rate was recorded during every session and exercise logs were completed to monitor adherence.	Continued standard care.	None reported
Hallsworth et al. (2015) <i>RCT</i>	HIIT	12 weeks	3 times per week	Sessions consisted of 5 intervals on a cycle ergometer at an intensity equating to 'very hard' (16-17) on a Borg RPE scale. Interval length progressed by 10 seconds per week from 2 minutes in week 1 to 3 minutes 50 seconds by week 12. Intervals were interspersed with 3 minutes consisting of 90 seconds passive recovery, 60 seconds upper body resistance band exercise and 30 seconds preparation for the subsequent interval.	The first 2 exercise sessions were supervised by members of the research team after which instructions were provided via voice-recordings loaded onto an iPod (Apple, CA, USA). Exercise diaries were completed and reported completion of 33 out of 36 prescribed sessions was considered 'adequate adherence'.	Continued standard care.	Participants were instructed to continue their normal routine and medical care, making no changes to diet, habitual activity or medication. Participants were asked to monitor and maintain body weight within 1% of baseline.

Haus et al. (2013) <i>Uncontrolled Intervention</i>	Aerobic	1 week	7 consecutive days	Sessions lasted 50-60 minutes consisting of 40-50 minutes of walking or running on a treadmill at 80-85% HR _{max} with appropriate warm-up and cool-down.	All exercise sessions were supervised by an exercise physiologist.	N/A	Participants were instructed to maintain normal dietary habits and habitual physical activity.
Hickman et al. (2013) <i>RCT</i>	Resistance	24 weeks	3 times per week	Circuit-based sessions consisting of 15 moderate-intensity (50% 1RM) resistance exercise covering the main muscle groups. Each exercise was performed for 30 seconds with 30 seconds rest, during which participants moved to the next exercise. Sessions consisted of 1 circuit (12 min) during week 1 and progressed to 5 circuits (60 min) by week 12. 1RM was re-assessed every 4 weeks.	All exercise sessions were supervised	NA	There were no prescribed dietary changes for the exercise intervention group.
Houghton et al. (2017) <i>RCT</i>	Combined (HIIT + Resistance)	12 weeks	3 times per week	A short HIIT session was performed made up of a 5 minute warm-up followed by 3 x 2 minute intervals at an intensity equating to 'very hard' (16-18) on a Borg RPE scale. Intervals were interspersed with 1 minute rest. This was immediately followed by a resistance exercise circuit that comprised of 5 whole-body exercise targeting large muscle groups. Participants lifted a weight that equated to an RPE of 14-16 ('hard').	All exercise sessions were supervised by an accredited exercise specialist and recorded to ensure adherence.	Continued standard care with maintenance of baseline weight.	Participants were instructed to maintain normal dietary habits and habitual physical activity.
Johnson et al. (2009) <i>RCT</i>	Aerobic	4 weeks	3 times per week	Each session lasted 30-45 minutes consisting of 15 minute bouts of cycling with 5 minute recovery periods. Intensity was increased from 50% of pre-training $\dot{V}O_2$ peak during week 1 to 60% in week 2 and 70% in weeks 3 and 4.	All exercise sessions were supervised	30 minute home-based whole-body stretching routine.	Participants were instructed to maintain habitual diet throughout the study. 24 hour food records were collected on the first and final three training sessions.

Keating et al. (2015) <i>RCT</i>	Aerobic	8 weeks	Ex1 (HI:LO): 3 sessions per week Ex2 (LO:HI): 4 sessions per week	Ex1 (HI:LO): 2 laboratory-based cycling sessions and 1 home-based brisk walking session per week all at the same intensity and duration. The programme progressed as follows: Week 1: 45 minutes at 50% VO ₂ peak. Week 2: Individual progression. Weeks 3-8: 60 minutes at 50% VO ₂ peak. Ex2 (LO:HI): 3 laboratory-based cycling sessions and 1 home-based brisk walking session per week all at the same intensity and duration. The programme progressed as follows: Week 1: 30 minutes at 50% VO ₂ peak. Week 2: Individual progression. Weeks 3-8: 45 minutes at 70% VO ₂ peak.	All laboratory exercise sessions were supervised by an accredited exercise physiologist. Adherence was 90, 96 and 94% in the HI:LO, LO:HI and control groups respectively.	Stretching, self-massage and 'fitball' programme. 1 session per week were performed in the laboratory and the remaining 2 at home. During the one supervised laboratory session participants performed 5 min of cycling 30 W to maintain ergometer familiarity.	Participants were instructed to maintain habitual diet throughout the study. Participants completed 24 hour food records on 3 non-exercise days at baseline and during the final week of the intervention Participants also wore a tri-axial accelerometer for 2 weeks before and after the intervention.
Langley et al. (2016) <i>Uncontrolled Intervention</i>	Combined (Aerobic + HIIT + Resistance)	12 weeks	3 times per week	Participants completed two whole-body strength training sessions and two sessions on a spinning bike per week. One bike session consisted of aerobic intervals for seven minutes at 85% HR _{max} with three minutes rest of active recovery against a light load between intervals. Participants completed three intervals in week one, four in weeks two-to-five and five from week six onwards. The second session consisted of two minute intervals at > 90% HR _{max} with two minutes rest of active recovery against a light load between intervals. Participants completed six intervals in week one, seven in weeks two-to-five and ten from week six onwards.	All exercise sessions were supervised Mean attendance was 90%.	N/A	Participants recorded habitual diet before and after the intervention.
Lee et al. (2013) <i>RCT</i>	Ex1: Aerobic Ex2: Resistance	13 weeks	3 times per week	Ex1: Aerobic exercise were performed on treadmill and/or elliptical. The programme progressed as follows: Week 1: 40 minutes at HR equating to ~50% VO ₂ peak. Weeks 2-8: 60 minutes at HR equating to ~70% VO ₂ peak. Ex2: 10 whole-body exercises targeting large muscle groups. All sessions were 60 minutes in duration. The programme progressed as follows: Weeks 1-4: 1-2 sets of 8-10 reps at 60% 1RM. Weeks 4-13: 2 sets of 8-12 reps to fatigue.	All exercise sessions were supervised by exercise physiology graduates. 4 participants did not complete exercise training (2 from each group). Mean (± SD) attendance was 95% (± 4.3%) and 97% (± 2.8%) in aerobic and resistance groups respectively.	Asked not to participate in any structured exercise activity To aid adherence, participants in control group were offered the opportunity to complete either exercise intervention following post-study assessment.	Participants consumed a weight-maintenance diet throughout the duration of the study.

Malin et al. (2013) <i>Uncontrolled Intervention</i>	Aerobic	1 week	7 consecutive days	Sessions lasted for approximately 60 minutes consisting of treadmill running at 85% HR _{max} .	All exercise sessions were supervised and 100% adherence was reported.	N/A	Participants were instructed to maintain normal dietary habits and habitual physical activity.
Meex et al. (2010) <i>Uncontrolled Intervention</i>	Combined (Moderate Aerobic + Resistance)	12 weeks	3 times per week total (2 aerobic, 1 resistance)	Aerobic: 30 minutes at 55% W _{max} . Resistance: 8 whole-body exercises targeting large muscle groups with 2 sets of 8 reps at 55% MVC.	Training sessions were supervised with 4 participants exercising per session.	N/A	None reported
Oh et al. (2014) <i>Uncontrolled Intervention</i>	Vibration / Acceleration	6 weeks	3 times per week	Whole-body exercises were performed using a vertical vibration machine. Sessions lasted 40 minutes with 30 seconds in between each exercise. Each week, one 'movement preparation', one 'strength and power' and one 'massage' session were performed.	Trained staff supervised all exercise sessions to ensure correct execution.	N/A	Participants received lifestyle counselling regarding diet and physical activity for NAFLD for 12 weeks prior to the intervention. This ceased at the beginning of the intervention. Participants completed 24 hour food records for 3 consecutive days and wore a uniaxial accelerometer for 2 weeks at baseline and during the final week of the intervention.
Pugh et al. (2014) <i>RCT</i>	Aerobic	16 weeks	3 – 5 times per week	Sessions consisted of a combination of treadmill and cycling exercise and progressed as follows: Weeks 1-4: 30 minutes at 30% HRR, 3 times per week. Weeks 4-8: 30 minutes at 45% HRR, 3 times per week. Weeks 8-12: 45 minutes at 45% HRR, 3 times per week. Weeks 12-16: 45 minutes at 60% HRR, 5 times per week.	Weekly sessions were supervised by a trained exercise physiologist.	Conventional care consisting of lifestyle advice from a consultant hepatologist or specialist nurse provided at a clinical consultation. No supervision or guidance was provided beyond this initial visit.	Participants in the exercise group were instructed to make no dietary modifications throughout the duration of the study, confirmed by 24 hour food diaries completed for 3 days before and after the intervention.
Sargeant et al. (2018) <i>Controlled Longitudinal Intervention</i>	HIIT	6 weeks	3 times per week	Sessions consisted of 30-second maximal sprints on cycle ergometer interspersed with 4.5-minute periods of active recovery at 50W. In addition participants completed a five-minute warm up and three-minute cool-down at 50W. Participants completed four intervals per session for the first two weeks after which an additional interval was added every two weeks, such that six intervals were completed per session in weeks five and six.	All exercise sessions were supervised and session attendance was 100%.	Each participant completed a six-week 'control phase' before the start of the training intervention, before and after which all study were assessed.	Participants were instructed to maintain habitual diet and lifestyle throughout the duration of the study.

Shojaee-Moradie et al. (2007) <i>RCT</i>	Aerobic (Vigorous)	6 weeks	3 times per week	Exercise performed at 60-85% VO ₂ max for a minimum of 20 minutes,	1 session per week was supervised by an exercise physiologist.	Participants were asked to continue their normal diet and lifestyle habits.	Participants were instructed not to change dietary habits throughout the study.
Sullivan et al. (2012) <i>RCT</i>	Aerobic	16 weeks	5 times per week	Sessions consisted of treadmill walking. During weeks 1-4, participants exercised for 15-30 minutes at 45-55% of pre-training VO ₂ peak. Sessions progressed regularly until participants performed 30-60 minutes of exercise at 45-55% VO ₂ peak and this was maintained for the remainder of the intervention.	1 session per week was performed under direct supervision from a member of the research team. The remaining sessions were completed at home.	Participants were instructed to maintain current activities of daily living and were contacted once per week to review compliance.	None reported
van der Heijden et al. (2010a) [‡] <i>Uncontrolled Intervention</i>	Aerobic	12 weeks	4 times per week	Sessions consisted of treadmill, elliptical or cycle ergometer exercise. Sessions lasted approximately 50 minutes consisting of 10 minutes warm-up, 10 minutes cool down and 30 minutes exercise at a heart rate corresponding to 70% of baseline VO ₂ peak (mean ± SE: 86 ± 2% HR _{max}).	2 exercise sessions per week were supervised by an exercise physiologist and the remaining were performed at home with adherence monitored by recording of heart rate. On average, participants completed 91 ± 2% of prescribed sessions.	N/A	Participants were instructed to make no changes to dietary or physical activity habits during the duration of the study. Body weight was monitored two times per week to assure weight stability.
van der Heijden et al. (2010b) <i>Uncontrolled Intervention</i>	Resistance	12 weeks	2 times per week	Sessions lasted 60 minutes and consisted of 10 whole-body resistance exercises targeting large muscle groups. The programme progressed as follows: Weeks 1-2: 2-3 sets of 8-12 reps at 50% of 3RM. Weeks 3-8: individual progression increasing firstly by number of reps followed by weight. Weeks 9-12: 3 sets of 15-20 reps at 80-85% 3RM.	All exercise sessions per week were supervised by an exercise physiologist. On average, participants completed 96 ± 1% of the prescribed sessions.	N/A	Participants were instructed to make no changes to dietary or physical activity habits during the duration of the study. Body weight was monitored two times per week to assure weight stability.
Zhang et al. (2016) <i>RCT</i>	Aerobic	24 weeks	5 times per week	Sessions lasted 30 minutes. In the moderate exercise group, participants walked briskly at approximately 120 steps per minutes so that their heart rate was 45 to 55% of predicted HR _{max} . In the vigorous exercise group, participants jogged at an intensity that elicited 65 to 80% of their predicted HR _{max} .	Participants were supervised during the first two to four weeks of training to familiarise themselves with the correct exercise intensity. Participant also attended bi-weekly health education sessions. Participants then performed sessions at a local community centre and received twice-weekly telephone calls to assess adherence. Participants in the moderate group were given pedometers to monitor training step rate.	Participants were instructed to maintain physical activity habits and attended bi-weekly health education sessions that were held separately to those attended by the intervention groups.	Participants were instructed to make no changes to their diet throughout the duration of the study.

*Study did not include 'standard care' or no 'intervention group'. Patients were randomised to either exercise or dietary interventions; ‡ Manuscript refers to same study as van der Heijden et al (2009); 1RM: 1-repetition maximum; 3RM: 3-repetition maximum; HIIT: High-intensity interval training; HR_{max}: Maximal heart rate; HRR: Heart rate reserve; NAFLD: Non-alcoholic fatty liver disease; N/A: Not applicable; RCT: Randomised controlled trial; reps: repetitions; RPE: Rating of perceived exertion; $\dot{V}O_2$ peak: peak oxygen consumption.

Table S4 – Risk of bias assessment for all studies

First Author (Publication Year)	Reference Number in Main Manuscript	Item Category / Number																										
		Reporting										External Validity			Internal Validity: Bias						Internal Validity: Confounding						Power	
		1	2	3	4	5	6	7	8	9	10	11	12	13	*15	16	17	18	19	20	21	22	23	24	25	26	27	
Cassidy (2016)	(ref. 34)	1	1	1	1	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	
Cuthbertson (2016)	(ref. 35)	1	1	1	1	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0	0	1	2	
Hallsworth (2011)	(ref. 36)	1	1	1	1	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0	0	1	2	
Hallsworth (2015)	(ref. 47)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0	1	2	
Haus (2013)	(ref. 39)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	0	0	0	0	1	0	
Hickman (2013)	(ref. 18)	1	1	1	1	2	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	2	
Houghton (2017)	(ref. 48)	1	1	1	1	2	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	0	1	0	0	1	2	
Johnson (2009)	(ref. 11)	1	1	1	1	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0	1	1	0	
Keating (2015)	(ref. 41)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	0	0	1	2	
Langley (2016)	(ref. 37)	1	1	1	1	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	
Lee (2013)	(ref. 43)	0	1	1	1	2	1	1	0	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	1	1	0	
Malin (2013)	(ref. 46)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	0	0	1	0	
Meex (2010)	(ref. 32)	0	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	0	1	1	0	
Oh (2014)	(ref. 42)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	0	0	0	1	0	
Pugh (2014)	(ref. 40)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	0	1	0	0	1	1	
Sargeant (2018)	(ref. 33)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	0	0	1	0	
Shojaee-Moradie (2007)	(ref. 49)	0	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0	1	0	
Sullivan (2012)	(ref. 12)	1	1	1	1	2	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	0	1	0	0	1	2	
van der Heijden (2009 & 2010a) [‡]	(refs. 51 & 44)	1	1	1	1	2	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	0	0	1	0	
van der Heijden (2010b)	(ref. 45)	1	1	1	1	2	1	1	0	1	0	0	0	1	0	0	1	1	1	1	1	0	0	0	1	1	2	
Zhang (2016)	(ref. 38)	1	1	1	1	2	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	2	

Full marking criteria can be found in the appendix of the original article by Downs and Black (13). * Item 14 of the Downs and Black scale (concerning participant blinding) was not scored in this meta-analysis as blinding participants to group allocation in exercise studies is very difficult and uncommon.

Item 27 (concerning power) was amended as outlined above in this electronic supplementary material. A higher number represents a more positive score. ‡ Two manuscripts reporting separate outcomes of the same study.

Table S5 - Pooled participant characteristics of RCTs included in meta-analysis of the effects of exercise training on IHTG (n = 10 studies)

	Exercise Groups	Control Groups
Total participant number	283	169
Male/Female/NR	111/149/23	77/77/15
Age (years)	52 (45 to 61)	52 (39 to 62)
Body weight (kg)	82.7 (71.4 to 103.0)	84.0 (72.1 to 113.7)
BMI (kg•m ⁻²)	30.2 (28.0 to 37.1)	30.2 (28.0 to 40.0)
Body fat (%)	34.4 (30.4 to 38.9)	34.8 (31.0 to 42.5)
Waist circumference (cm)	99.9 (95.5 to 110.0)	99.2 (93.7 to 109.0)
Fasted glucose (mmol•L ⁻¹)	5.63 (4.30 to 6.80)	5.67 (4.00 to 7.00)
$\dot{V}O_2$ peak (ml•kg ⁻¹ •min ⁻¹)	23.6 (21.8 to 26.4)	22.3 (18.5 to 27.0)
IHTG (%)	16.2 (6.9 to 21.3)	14.5 (7.1 to 21.4)

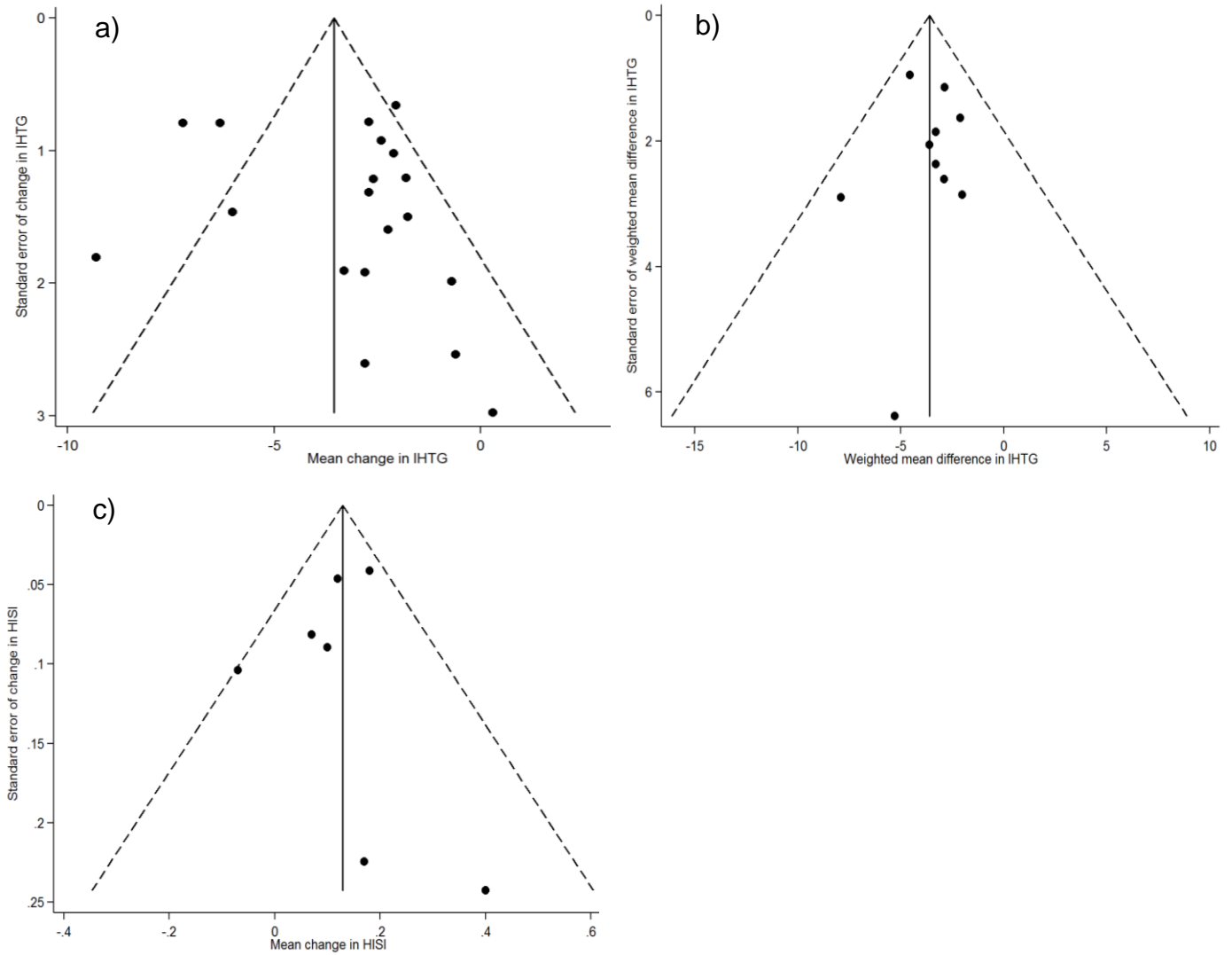
Data presented as weighted mean with range; BMI: body mass index; IHTG: intrahepatic triglyceride; NR: not reported; RCTs: randomised controlled trials; $\dot{V}O_2$ peak: peak oxygen uptake.

Table S6 – Pooled intervention characteristics of RCTs included in meta-analysis of the effects of exercise training on IHTG (n = 10 studies)

	Mode	Range
Intervention duration (weeks)	12	4 to 24
Session frequency (times per week)	3	3 to 5
Session duration (mins)	-	30 to 53

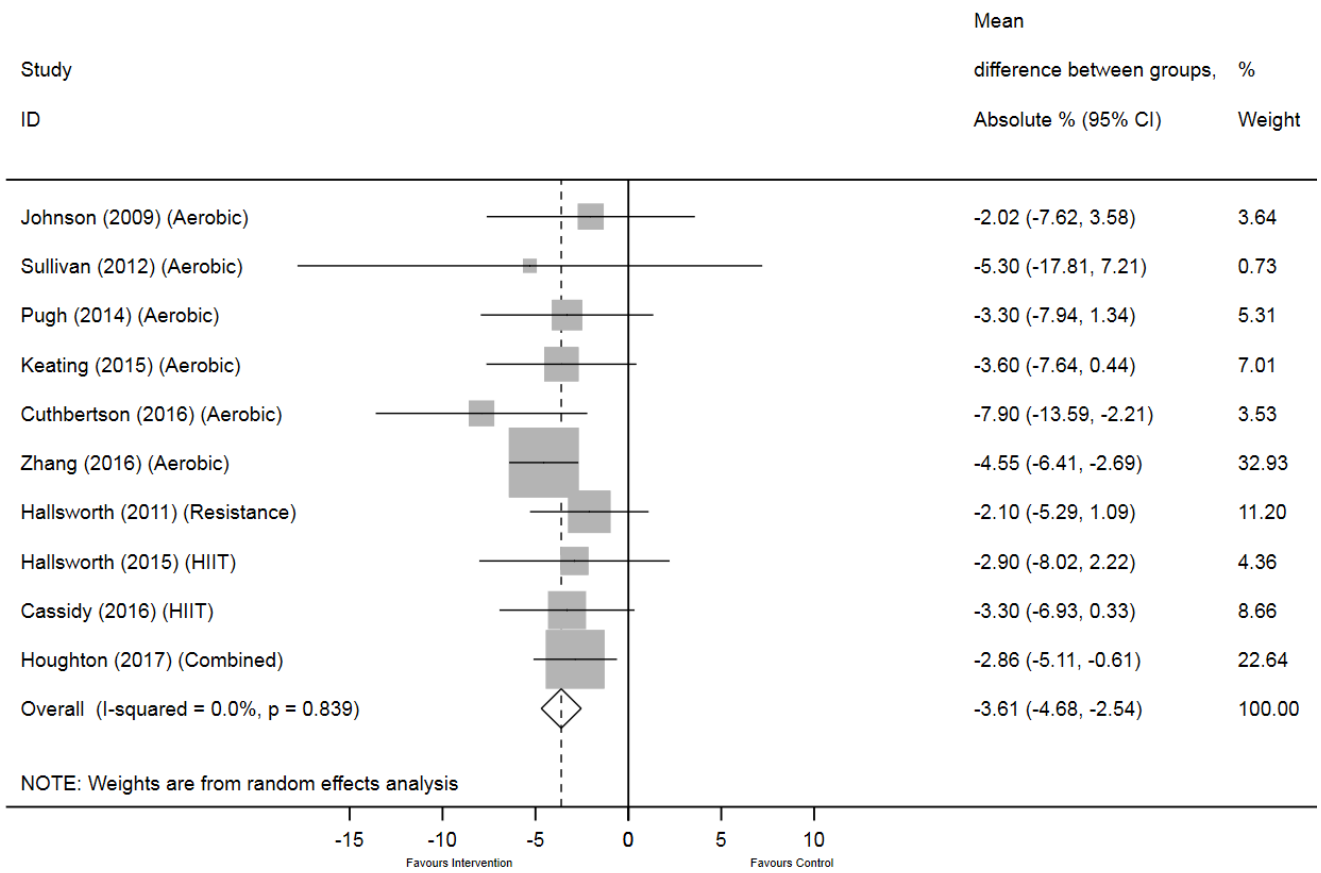
IHTG: intrahepatic triglyceride; NR: not reported; RCTs: randomised controlled trials.

Figure S1 – Funnel plots for assessment of publication bias in studies included in the following analyses: a) within-group change in IHTG, b) difference between groups in change in IHTG, and c) within-group change in HISI.



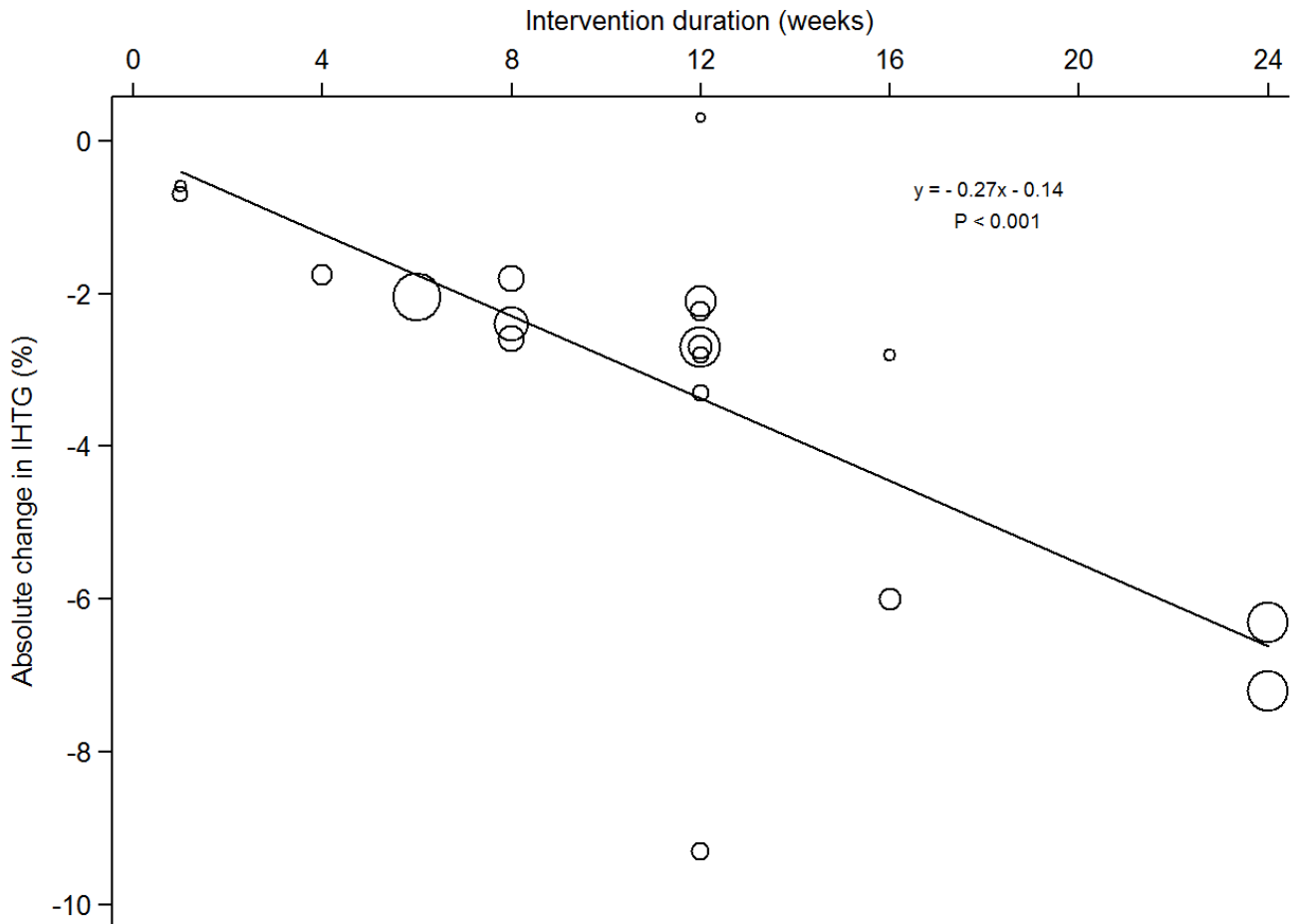
HISI: hepatic insulin sensitivity index; IHTG: intrahepatic triglyceride

Figure S2 – Meta-analysis of the pooled effect of exercise training on IHTG using eligible RCTs only.



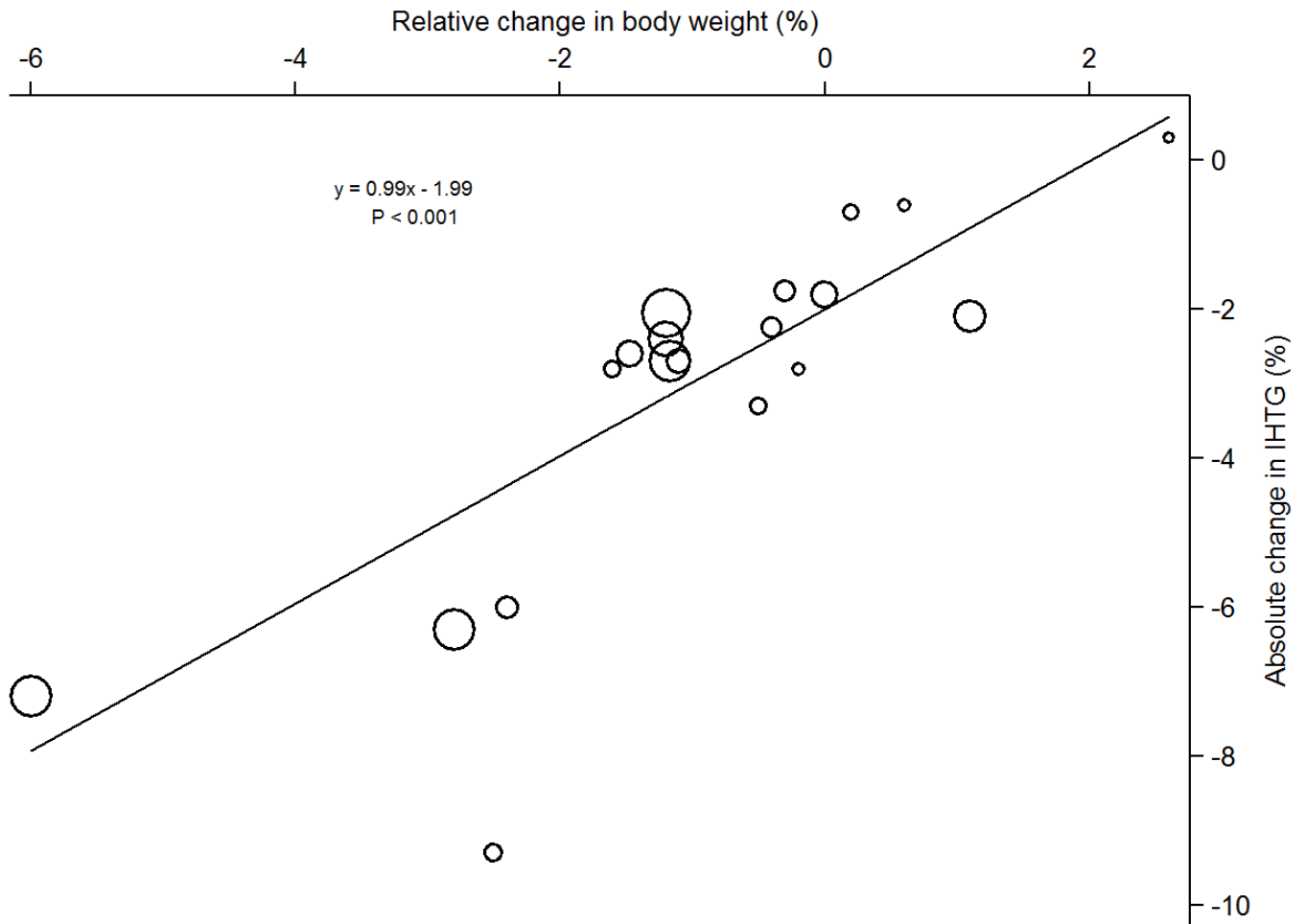
95% CI: 95% confidence interval; Combined: combined HIIT-plus-resistance exercise; HIIT: high-intensity interval training; IHTG: intrahepatic triglyceride; RCTs: randomised controlled trials.

Figure S3 – Meta-regression between duration of exercise intervention in weeks and the absolute change in IHTG from baseline elicited in all eligible studies.



Each circle represents a study, with larger circles indicating greater weight within the meta-regression analyses. IHTG: intrahepatic triglyceride.

Figure S4 – Meta-regression between the magnitude of weight loss, relative to baseline, elicited by exercise intervention and the absolute change in IHTG



Each circle represents a study, with larger circles indicating greater weight within the meta-regression analyses. IHTG: intrahepatic triglyceride.

Bibliography

1. Higgins J, Green S. *Cochrance Handbook for Systematic Reviews of Interventions*. Chichester: John Wiley & Sons Ltd.; 2011.
2. Pugh CJA, Sprung VS, Kemp GJ, et al. Exercise training reverses endothelial dysfunction in nonalcoholic fatty liver disease. *AJP Hear Circ Physiol* 2014; **307**: H1298–H1306.
3. Hallsworth K, Thoma C, Hollingsworth KG, et al. Modified high-intensity interval training reduces liver fat and improves cardiac function in non-alcoholic fatty liver disease: a randomized controlled trial. *Clin Sci* 2015; **129**: 1097–1105.
4. Houghton D, Thoma C, Hallsworth K, et al. Exercise reduces liver lipids and visceral adiposity in patients with non-alcoholic steatohepatitis in a randomized controlled trial. *Clin Gastroenterol Hepatol* 2017; **15**: 96–102.
5. Cuthbertson DJ, Shojaee-Moradie F, Sprung VS, et al. Dissociation between exercise-induced reduction in liver fat and changes in hepatic and peripheral glucose homoeostasis in obese patients with non-alcoholic fatty liver disease. *Clin Sci* 2016; **130**: 93–104.
6. Hickman IJ, Byrne NM, Croci I, et al. A Pilot Randomised Study of the Metabolic and Histological Effects of Exercise in Non-alcoholic Steatohepatitis. *J Diabetes Metab* 2013; **4**: 300.
7. Keating SE, Hackett DA, Parker HM, et al. Effect of aerobic exercise training dose on liver fat and visceral adiposity. *J Hepatol* 2015; **63**: 174–182.
8. Langleite TM, Jensen J, Norheim F, et al. Insulin sensitivity, body composition and

adipose depots following 12 w combined endurance and strength training in dysglycemic and normoglycemic sedentary men. *Arch Physiol Biochem* 2016; **122**: 167–179.

9. Meex RCR, Schrauwen-Hinderling VB, Moonen-Kornips E, et al. Restoration of muscle mitochondrial function and metabolic flexibility in type 2 diabetes by exercise training is paralleled by increased myocellular fat storage and improved insulin sensitivity. *Diabetes* 2010; **59**: 572–579.
10. Sargeant JA, Bawden SJ, Aithal GP, et al. Effects of sprint interval training on ectopic lipids and tissue-specific insulin sensitivity in men with non-alcoholic fatty liver disease. *Eur J Appl Physiol* 2018; **118**: 817-828.
11. van der Heijden G-J, Wang Z, Chu Z, et al. Strength exercise improves muscle mass and hepatic insulin sensitivity in obese youth. *Med Sci Sports Exerc* 2010; **42**: 1973–1980.
12. Zhang H-J, He J, Pan L-L, et al. Effects of moderate and vigorous exercise on nonalcoholic fatty liver disease. *JAMA Intern Med* 2016; **176**: 1074–1082.
13. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Heal* 1998; **52**: 377–384.