The effect of exercise training on liver function in adults who are overweight or exhibit fatty liver disease:  
A systematic review and Meta-Analysis.

Smart N.A. PhD¹, King N. PhD², McFarlane J.R. PhD¹, Graham P.L. PhD³, Dieberg G. PhD¹

Running Head: Exercise for fatty liver disease

1. School of Science and Technology, University of New England, Armidale, NSW 2350, Australia.
2. School of Biomedical and Healthcare Sciences, Plymouth University, Plymouth, PL4 8AA. UK
3. Department of Statistics, Macquarie University, Sydney, NSW 2109, Australia

Address for Correspondence:

Associate Professor Neil Smart  
School of Science and Technology  
University of New England  
Armidale, NSW 2351, Australia  
E-mail: nsmart2@une.edu.au  
Tel: +61 2 6773 4076  
Fax: +61 2 6772 8997

Word Count:

Abstract: 265  
Main Text: 4696
Keywords: exercise training, liver function, fatty liver disease

Author Contributions

Dr Nicola King is a post-doctoral lecturer who conducted some of the data extraction, study quality assessment and manuscript writing.

Dr Gudrun Dieberg assisted with data searching, extraction, data checking, study quality assessment and writing and editing of the manuscript.

Prof. Jim McFarlane assisted with the literature search, assessment of study quality and verified the accuracy of the main text and supplementary files section as well as contributing to writing and editing the main text.

Dr Petra Graham, verified accuracy of data extraction, conducted analyses especially regarding heterogeneity, provided comprehensive statistical advice for the response to reviewers as well as contributing to the editing of the rebuttal and final revised text and provision of some of the figures.

A/Prof. Smart was the team leader, who also assisted with study inclusion/exclusion, data extraction/analysis and writing and editing of both the main text and supplementary files.

This work was completed without funding and here are no competing interests or financial disclosures for any of the authors. Ethical approval was not required as this was a literature-based analysis and there are no limitations on data sharing.

Acknowledgements

None
Abstract

Objective Exercise training has been shown to have beneficial effects on liver function in adults overweight or with fatty liver disease. To establish which exercise program characteristics were likely to elicit optimal improvements.

Design: Systematic review and meta-analysis of randomized, controlled trials.


Eligibility criteria for selecting studies: exercise intervention, with or without dietary intervention, versus usual care in adults undertaking, exercise training, who were overweight, obese or exhibited fatty liver disease (NAFLD or NASH).

Results: We included 21 randomized controlled trials, totalling 1530 participants. Exercise intervention studies with total exercise program workload >10,000Kcal produced significant improvements in intrahepatic fat, -3.46% [95% CI -5.20, -1.73], p<0.0001, I²=73%; Effect Size (SMD) -1.77 [-3.11, -0.42], p=0.01, I²=77%.

When data from only exercise studies were pooled; there was a reduction in fasting free fatty acids (FFA) -74.15 µmol.L⁻¹ [95% CI -118.47, -29.84], p=0.001, I²=67% with a large effect size (SMD) -0.94 [-1.36, -0.52], p<0.0001, I²=0%. When data from only exercise studies were pooled; there was a significant reduction in insulin MD -1.88 UL [95% CI -3.43, -0.34], p=0.02, I²=31%. The liver enzymes Alanine aminotransferase (ALT), aspartate aminotransferase (AST) and Gamma-glutamyl transpeptidase (GGT) were not significantly altered with exercise.

Conclusions Exercise training reduces intrahepatic fat and free fatty acids while increasing cardiorespiratory fitness. An aggregate exercise program energy expenditure (>10,000 Kcal) may be required to promote reductions in intrahepatic fat.

Systematic review registration: Not registered.
SUMMARY

- Several new published randomized, controlled trials of exercise intervention suggest, for the first time, non-esterified fatty acids are reduced with exercise intervention
- Changes in intra-hepatic may be possible, but only in exercise programs that elicit >10,000Kcal of energy expenditure
- Varied study interventions render it difficult to separate the relative effects of diet and exercise on liver function

HOW MIGHT FINDINGS INFLUENCE CLINICAL PRACTICE

- Imaging, rather than liver enzyme concentrations, should be used to assess liver status
- Preferably 3 months of adherent exercise training, after baseline assessment, should ensue before liver imaging is repeated.
Introduction

Central obesity is linked to raised serum liver enzymes and future risk of developing insulin resistance[1]. Fatty liver disease (FLD) and non-alcoholic fatty liver disease (NAFLD) are also associated with increased serum levels of liver enzymes[2]. FLD may be diagnosed through liver function tests and imaging, such as magnetic resonance imaging or ultrasound of the liver, to detect excess intrahepatic fat. Exercise can greatly assist in reducing obesity and FLD, mainly by improving weight loss through dietary changes and exercise and reducing alcohol consumption[3].

Obesity and FLD are risk factors for diabetes and cardiovascular disease. The prevalence of FLD varies greatly from a 10–35% prevalence rate in the US with approximately 2–5% of patients with non-alcoholic steatohepatitis (NASH)[4], the more severe form of NAFLD.

A 2012 systematic review by Thoma et al[5] examined the effects of diet and exercise on liver function, but although data pooling was not performed, the authors concluded that dietary induced weight loss may offer greater liver health benefits than exercise. A 2012 systematic review, of randomized and non-randomized controlled studies in obese and FLD populations, by Keating and colleagues examined intrahepatic liver fat content and Alanine aminotransferase (ALT). The pooled analysis reported that the former, but not the latter, was reduced with exercise training of 4-26 weeks duration[6]. A number of new studies, including two studies with >100 participants[7, 8], have been published after Keating’s analysis. Moreover, Keating et al. pooled data from different modalities of exercise (aerobic and resistance) and did not adjust for the great variation in exercise program durations (1-12 months), frequency (2-7 sessions weekly) and intensity. There are now sufficient randomized controlled trials available to conduct analyses of several additional outcome measures and to adjust for variation in exercise program characteristics to identify those that are most beneficial in patients with obesity and fatty liver
disease. Furthermore liver enzymes may not be sensitive enough to be the only measure of liver function or liver fat[9], so new studies mean other markers can be examined.

The primary aim of this work was to conduct a systematic review and meta-analyses to establish the effect of exercise training on surrogate markers of liver function in adults who were overweight or exhibited fatty liver disease. The secondary aim was to establish if there is an exercise program volume more likely to elicit optimal improvements in related health outcome measures.
Materials and Methods

Search strategy

Potential studies were identified by conducting a systematic search using PubMed, www.ncbi.nlm.nih.gov/pubmed (1966 to October 2, 2015), the Pub Med search strategy can be seen in supplementary files. CINAHL and the Cochrane controlled trials registry were also searched (1966-October 2, 2015). The search strategy included the key concepts of overweight, obese, fatty liver disease, non-alcoholic fatty liver disease, non-alcoholic steatohepatitis, lifestyle therapy, physical training and exercise training. These were combined with a sensitive search strategy to identify randomized controlled trials (Supplementary Figure S1). Titles and abstracts were screened by NS and GD and full-text articles were downloaded for those meeting inclusion criteria. Reference lists of papers found were also scrutinized for new references. All identified papers were assessed independently by two reviewers (GD and NS), a third reviewer (JM) was consulted to resolve disputes. Searches of published papers were also conducted up until October 2, 2015.

Inclusions

Randomized controlled trials of progressive aerobic, resistance, or combined exercise training in adults (>18 years) who were overweight or obese or exhibited fatty liver disease (NAFLD or NASH) were included. The definition of obese/overweight was made by the individual studies, but this was verified by checking baseline data. Combined exercise training was defined as interventions that utilized both aerobic and resistance training simultaneously. We included studies comparing two types of exercise (e.g. aerobic versus resistance) or exercise versus usual care. Included studies employed three sessions or more of exercise, in order to differentiate from acute exercise effects. Studies involving dietary control or intervention groups were included only if the diet was the same between exercise and control groups allowing the independent effects of exercise to be examined. There were no language restrictions.
Exclusions

Animal studies, review papers and non-randomized controlled trials were excluded. Studies with unmatched intervention versus control group participants were excluded. We also excluded studies of healthy, normal body mass participants as liver function was likely to be normal. Authors were contacted to provide missing data or to clarify if data was duplicated in multiple publications. Incomplete data, or data from an already included study, were excluded. Studies using interventions other than exercise and/or (e.g. electroacupuncture, ultrasound) were excluded.

Data Extraction

Data on outcome measures were archived in a database, data were extracted by NS and verified by GD, JM was consulted if discrepancies occurred. The outcome measures were: change in intrahepatic fat (percentage change), Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Gamma-glutamyl Transpeptidase (GGT), free fatty acids (FFA), body mass and body mass index, insulin, total cholesterol and peak VO₂ Table 1. Volume – where possible energy expenditure for weekly and total exercise programs. We calculated energy expended by establishing oxygen consumption at the training intensity and using 5Kcal per litre of Oxygen consumed, thus establishing Kcals expended per minute. We then multiplied Kcals/min by session duration (mins) and then multiplied this by total number of sessions to calculate aggregate energy expenditure for each exercise program, this is an approach we have utilized previously [10, 11] [12]. We also conducted analysis of effects on interhepatic fat only in exercise programs that expended an aggregate of >10,000Kcals.

Data Synthesis

The meta package in R version 3.2.1 was used to complete the meta-analysis and generate forest plots [13, 14]. Pooled data are presented as mean differences (MD) and as the effect size Hedges’ g i.e. standardised mean differences (SMD) with 95% confidence interval for intervals with a significant MD. We chose a random effects model as we anticipated considerable
heterogeneity. A minimum of three studies was required for meta-analysis to be undertaken. Multivariate meta-analysis (using the metafor package)[15] with a random intercept for study was used for studies that compared more than one intervention group with a control.

Meta-analyses were completed for continuous data by using the change in the mean and standard deviation of outcome measures. It is an accepted practice to only use post intervention data for meta-analysis but this method assumes that random allocation of participants always creates intervention groups matched at baseline for age, disease severity etc. Change in post intervention mean was calculated by subtracting baseline from post intervention values. Data required was either (i) 95% confidence interval data for pre-post intervention change for each group or when this was unavailable (ii) actual p values for pre-post intervention change for each group or if only the level of statistical significance was available (iii) we used default p values e.g. p<0.05 becomes p=0.049, p<0.01 becomes P=0.0099 and p = not significant becomes p=0.05.

Effect sizes are interpreted in the usual way [16] with an effect size of 0.2 defined as small, 0.5 defined as moderate and 0.8 or higher defined as large.

Analyses

Where sufficient studies were included, meta-analysis was performed for the following groups (i) aerobic versus resistance training studies (ii) exercise only versus control studies and (iii) exercise plus diet versus diet control studies.

Several sub-analyses were conducted as follows:

In order to avoid considerable heterogeneity[17] we only presented forest plots when $I^2$ heterogeneity was <75%. In the exercise only interventions we conducted meta-analyses for:

Free fatty acids (FFA), Gamma-Glutamyl Transpeptidase (GGT), Body Mass Index (BMI), Insulin and Total Cholesterol.

Sensitivity Analysis
Heterogeneity

Heterogeneity was quantified using the $I^2$ test\cite{18}, as it does not inherently depend upon the number of studies considered. The $I^2$ statistic and corresponding 95% confidence interval (CI) were presented, to gauge the degree of heterogeneity present in sample \cite{18, 19}. $I^2$ values range from 0% (homogeneity) to 100% (greater heterogeneity); a CI that does not include 0% indicates that the hypothesis of homogeneity is rejected, and an inference of heterogeneity is merited\cite{18}.

Meta-analysis and forest plots are presented only for those studies in which the heterogeneity (as measured by $I^2$) was <75%. Heterogeneity beyond this level was deemed too heterogeneous\cite{17} to be sensibly pooled and while interest was in trying to identify factors responsible for the heterogeneity, none of the subsets contained the recommended minimum number of studies (10) to allow this.

Publication Bias

Egger bias tests (when at least 10 studies were included) and funnel plots\cite{20} were provided to assess the risk of publication bias (supplementary files).

Risk of Bias Assessment

Study quality was assessed by using the TESTEX scale\cite{21} - the Tool for the assEssment of Study qualiTy and reporting in EXercise (TESTEX) - a study quality and reporting assessment tool, designed specifically for use in exercise training studies. The main point of difference in TESTEX is that there are accommodations for: Activity monitoring in control groups to measure crossover to exercise by sedentary control patients; Assessment of the existence and method of activity monitoring in both exercise intervention and sedentary controls; Assessment of whether the relative exercise intensity remained constant and therefore potentially avoided de-training as participants initially adapt to new exercise programs; Assessment of whether periodic evidence-based adjustment of exercise intensity is reported exercise volume and exercise expenditure Information on all exercise characteristics (intensity, duration,
frequency and mode) is provided to calculate exercise volume and exercise energy expenditure.

This tool is a 15-point scale (5 points for study quality and 10 points for reporting) and addresses previously unmentioned quality assessment criteria specific to exercise training studies. Two reviewers NS and GD conducted the risk of bias assessment, JM was consulted of discrepancies occurred.
Results

Studies included in the review

Our initial search identified 53 manuscripts, hand searching of reference lists of included studies and key articles such as related reviews and the latest editions of relevant journals yielded a further 3 manuscripts. Out of 56 studies, 6 were excluded at first inspection as duplicates, 17 were not controlled trials of exercise therapy, 2 were excluded as they had participants <18 years, 2 were excluded as they were not randomized trials, 6 used unmatched interventions or comparator groups and 2 were counselling interventions encouraging exercise participation, leaving 21 included studies for analysis (PRISMA Statement - Figure 1).

Description of Included Studies

Our 21 included randomized controlled trials (25 intervention groups), had an aggregate of 1530 participants, 884 exercise participants and 646 controls. Table 2 summarizes included studies[7, 8, 22-40]. Supplementary Table S1 in the supplementary files summarizes the excluded controlled trials[41-52].

Thirteen studies used only an exercise intervention and eight studies used both exercise and dietary interventions. Five studies utilized resistance exercise and 15 utilized aerobic exercise, 4 studies also used combined aerobic and resistance exercise. Study duration ranged from 4 to 52 weeks, training frequency ranged from 2-5 times weekly with 20-60 minutes session duration. Aerobic training intensity ranged from 45-85% of peak VO₂. Control groups were classified as sedentary, although some used a stretching routine.

Measurement of Study Outcomes

Of note is that only one of the included studies[40] used the gold standard liver biopsy method to assess liver function. Instead of direct assessment of liver
function several studies used plasma concentrations of liver enzyme as surrogate markers of liver function. Moreover a number of different techniques were used to quantify liver enzymes.

**Study Quality Assessment**

Median TESTEX score was assessed as 10 out of 15 by two reviewers NS and GD, three studies each scored 7, 8, or 11. Four studies scored 9 and five studies scored 10, one study scored 12, two studies scored 13 (Figure 2). Of the TESTEX items the following were done particularly poorly in general; Allocation concealment only 5/21 studies; Blinding of Assessors 5/21; Physical activity monitoring in the control groups to check with sedentary controls crossed over to exercise 10/21; Assessment of energy expended during exercise 10/21.

**Summary of Change in Outcome Measures**

Only two included studies compared aerobic versus resistance training, as such this training group is not considered further.

**Change in Intrahepatic Fat**

Data on interhepatic fat (IHF) were available from 9 studies (12 intervention groups) of exercise, with or without dietary intervention. Six study groups showed significant IHF reductions, one study showed a significant increase and 5 studies showed no change.

Exercise only studies had high heterogeneity ($I^2$: 84.4% 95% CI 70.1%, 91.4%) and there were too few studies (6) to allow for investigation of the cause of the heterogeneity beyond the following subgroup analysis.

When data from the exercise only versus control programs that expended an aggregate of $>10,000$Kcal were pooled; there was a reduction in interhepatic fat $-3.46\%$ [95% CI -5.20, -1.73], $p<0.0001$, $I^2=73\%$ [95% CI 9%, 92%], sFigure 3. The effect size was large (SMD) $-1.77$ [-3.11, -0.42], $p=0.01$, $I^2=77\%$.

Diet plus exercise studies had no evidence of heterogeneity ($I^2$:0 95% CI 0%, 84.2%) and indicated no evidence of a change in interhepatic fat between diet and exercise versus diet alone studies MD $0.87\%$ [95% CI -0.07, 1.81], $p=0.07$ (from a multivariate model). Effect size (SMD) was small to moderate $0.29$ 95% CI: [-0.19, 0.78], $p= 0.24$
Change in Body Mass and Body Mass Index (BMI)

Fourteen exercise studies with or without dietary intervention (17 intervention groups) reported change in body mass. Nine groups reported a significant fall in body mass, 1 group reported a significant rise and 7 groups were equivocal.

The exercise only body mass studies (n=7) exhibited substantial heterogeneity (I^2 95.1% 95% CI 92.6%, 96.8%) even in the multivariate analysis. The small number of studies precluded investigation of the heterogeneity. The diet and exercise studies showed less heterogeneity (I^2 71.3% 95% CI 40.8% to 86.1%) but no evidence of a change in body mass in the multivariate meta-analysis model MD -2.94 Kg (95% CI -5.93, 0.04), effect size (SMD) was moderate -0.34 95% CI: (-0.55, -0.13) P= 0.001.

Eleven exercise study groups, with (n=5) or without (n=6) dietary intervention, reported change in BMI, with only 4 groups reporting a significant improvement, while the other 7 groups showed no change.

Meta-analysis of the 5 diet and exercise BMI studies indicated substantial heterogeneity: I^2 83% (95% CI 62%, 93%). The small number of studies precluded investigation of the heterogeneity. Meta-analysis of the effect of exercise only intervention studies on BMI found a no evidence of a change, MD -0.05 Kg.m^-2 [95%CI -0.16, 0.06] p=0.38, I^2=37% (95% CI 0%, 75%), see Figure 4. The effect size was small to moderate; (SMD) -0.26 [-0.56, 0.04], p=0.09, I^2=0%, and also indicated no change in BMI.

Fasting free fatty acids (FFA)

When data from all 9 exercise studies with or without dietary intervention (11 intervention groups) were examined, only 5 groups reporting a significant improvement, while the other 6 groups showed no change.

When data from only exercise studies were pooled using a multivariate meta-analysis model; there was a reduction in fasting FFA -74.15 µmol.L^-1 [95% CI -118.47, -29.84], p=0.001, I^2=67% (95% CI 32%, 85%), Figure 5. The effect size (SMD) indicated a large change in FFA -0.94 [-1.36, -0.52], p<0.0001, I^2=0%. The 3
pooled diet and exercise studies exhibited substantial heterogeneity; I² 88% 95% CI (65%, 96%) and were not considered further.

Liver Enzymes

Alanine aminotransferase (ALT)

Examination of data from 16 exercise studies with or without dietary intervention (20 intervention groups) revealed that ALT was not significantly altered in 10 groups and was significantly reduced (improved) in 5 groups and increased in 5 groups.

The 9 pooled exercise only studies exhibited substantial heterogeneity (I² 91% 95% CI 87%, 94%) and were not considered further. The 7 pooled diet and exercise studies exhibited high levels of heterogeneity (I² 73%, 95% CI 45%, 87%) with no evidence of a difference in effect between diet and exercise and diet control groups MD -0.96 IU.L⁻¹ (95% CI -2.84, 4.76). Effect size (SMD) negligible 0.01 95% CI: (-0.40, 0.42). P = 0.95.

Aspartate aminotransferase (AST)

When data from 9 exercise studies with or without dietary intervention (12 intervention groups) were examined, there was no change in AST in 6 groups, while 3 groups showed a significant reduction and 3 groups showed a significant increase.

The 4 pooled exercise only studies exhibited substantial heterogeneity (I² 88% 95% CI 77%, 94%) and were not considered further. The 5 pooled diet and exercise studies exhibited high levels of heterogeneity (I² 61%, 95% CI 5%, 84%) with no evidence of a difference in effect between diet and exercise and diet control groups MD -0.68 IU.L⁻¹ (95% CI -2.54, 1.18). Effect size (SMD) was small -0.20 95% CI: (-0.73, 0.34). P= 0.47.

Gamma-Glutamyl Transpeptidase (GGT)

Data from 6 exercise studies with or without dietary intervention (7 intervention groups) showed there was no significant change in GGT in 6 groups, only one
group showed a significant reduction. Only two diet and exercise studies were included thus pooling was not performed for that group.

When data from the 4 exercise only studies were pooled using a multivariate meta-analysis model; there was no significant reduction in GGT MD -3.52 IU.L⁻¹ [95%CI -8.37, 1.34] p=0.16, I²=71% (95% CI 26%, 89%), see Figure 6. The effect size (SMD) was moderate but the interval suggested no significant effect -0.30 [-0.69, 0.09], p=0.13, I²=74%.

**Peak VO₂**

Data from 8 exercise studies with or without dietary intervention (11 intervention groups) were examined, there was a significant improvement in cardio-respiratory fitness in 6 groups but not in 5 groups.

The 5 pooled exercise only studies exhibited substantial heterogeneity (I² 86% 95% CI 75%, 92%) and were not considered further. The 3 pooled diet and exercise studies exhibited high levels of heterogeneity (I² 68%, 95% CI 8%, 89%) with no evidence of a difference in effect between diet and exercise and diet control groups MD -0.05 ml O₂.kg⁻¹.min⁻¹ (95% CI -0.22, 0.12). Effect size (SMD) was small 0.1430 95% CI: (-0.5377, 0.8236). P= 0.68.

**Insulin**

Data from 11 exercise studies (12 intervention groups) with or without dietary intervention was analysed, there was a significant insulin reduction in 5 groups but no change in 6 groups, 1 group showed an increase in insulin.

When data from the 4 exercise only studies were pooled; there was a significant reduction in insulin MD -1.88 IU.L⁻¹ [95% CI -3.43, -0.34], p=0.02, I²=31% (95% CI 0%, 75%), Figure 7. The effect size was moderate but was not significant SMD -0.47 [-1.07, 0.13], p=0.13, I²=63%. Pooling the 7 diet and exercise studies resulted in substantial heterogeneity, I² 78% [95% CI 57%, 89%] hence these results are not considered further.

**Total Cholesterol**
Data from 12 exercise studies (12 intervention groups) with or without dietary intervention were considered, there was a significant total cholesterol reduction in 6 groups but no change in 7 groups.

When data from the 6 exercise only studies were pooled; there was a significant reduction in total cholesterol MD $-7.04\text{ mg.dL}^{-1}$ [-11.96, -2.13], $p=0.005$, $I^2=44\%$; (95% CI 0%, 78%), Figure 8 (top). The effect size was moderate; SMD $-0.49$ [-0.95, -0.03], $p=0.035$, $I^2=50\%$.

When data from the 6 diet and exercise studies were pooled using a multivariate meta-analysis model; there was a significant reduction in total cholesterol MD $-2.47\text{ mg.dL}^{-1}$ [-4.55, -0.39], $p=0.020$, $I^2=44\%$; (95% CI 0%, 78%), Figure 8 (bottom). The effect size was moderate; SMD $-0.49$ [-0.85, -0.13], $p=0.008$, $I^2=52\%$. 
Table 1. List of outcome measures and other data extracted from included studies.

<table>
<thead>
<tr>
<th>Data</th>
<th>Type</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-Hepatic Fat</td>
<td>Outcome Measure</td>
<td>% of interhepatic fat</td>
</tr>
<tr>
<td>Body Mass</td>
<td>Outcome Measure</td>
<td>Mass in Kg</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>Outcome Measure</td>
<td>Kg.m⁻²</td>
</tr>
<tr>
<td>Free Fatty Acids</td>
<td>Outcome Measure</td>
<td>µmol.L⁻¹</td>
</tr>
<tr>
<td>Gamma-Glutamyl Transpeptidase</td>
<td>Outcome Measure</td>
<td>Serum IU.L⁻¹</td>
</tr>
<tr>
<td>Alanine Aminotransferase</td>
<td>Outcome Measure</td>
<td>Serum IU.L⁻¹</td>
</tr>
<tr>
<td>Aspartate aminotransferase</td>
<td>Outcome Measure</td>
<td>Serum IU.L⁻¹</td>
</tr>
<tr>
<td>Peak VO₂</td>
<td>Outcome Measure</td>
<td>ml O₂.kg⁻¹.min⁻¹</td>
</tr>
<tr>
<td>Insulin</td>
<td>Outcome Measure</td>
<td>Serum IU.L⁻¹</td>
</tr>
<tr>
<td>Total Cholesterol</td>
<td>Outcome Measure</td>
<td>Serum mg./dL⁻¹</td>
</tr>
<tr>
<td>Session Duration</td>
<td>Co-variate</td>
<td>Minutes</td>
</tr>
<tr>
<td>Program Duration (Weeks)</td>
<td>Co-variate</td>
<td>Weeks</td>
</tr>
<tr>
<td>Session Frequency</td>
<td>Co-variate</td>
<td>Sessions Per Week</td>
</tr>
<tr>
<td>Exercise Mode</td>
<td>Co-variate</td>
<td>Type of Activity</td>
</tr>
<tr>
<td>Exercise Intensity</td>
<td>Co-variate</td>
<td>% Maximum Heart Rate</td>
</tr>
<tr>
<td>Year</td>
<td>Year of publication</td>
<td>Year</td>
</tr>
</tbody>
</table>
Table 2. Included Studies – Exercise Programming Details

<table>
<thead>
<tr>
<th>STUDY</th>
<th>Study Duration, Frequency &amp; Session Times</th>
<th>Intervention</th>
<th>Exercise Intensity</th>
<th>Total Exercise Program Energy Expenditure (Kcal)</th>
<th>No. of Subjects Exercise (control)</th>
<th>Method of IHF Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise vs control or exercise vs exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al-Jiffri 2013</td>
<td>12 wks 3 times/wk 30 mins</td>
<td>A v Control</td>
<td>A: 65-75% Max HR</td>
<td>NR</td>
<td>A:50 C:50</td>
<td>Liver Biopsy</td>
</tr>
<tr>
<td>Bacchi 2013</td>
<td>16 wks 3 times/wk 60 mins</td>
<td>A vs R</td>
<td>A: 60-65% HRR R: 70-80% IRM</td>
<td>NR</td>
<td>A: 13 R: 17</td>
<td>MRI</td>
</tr>
<tr>
<td>Balducci 2014</td>
<td>52 wks 2 times/wk NR</td>
<td>C vs Control</td>
<td>55-75% peak VO₂ 60% 1RM</td>
<td>NR</td>
<td>C: 303(303)</td>
<td>NR</td>
</tr>
<tr>
<td>Bonekamp 2008</td>
<td>26 wks 3 times/wk 45 mins</td>
<td>C vs Control</td>
<td>A: moderate R: 50% 1RM</td>
<td>13,650</td>
<td>28 (17)</td>
<td>MRS</td>
</tr>
<tr>
<td>Hallsworth 2011</td>
<td>8 wks 3 times/wk 45-60 mins</td>
<td>R vs Control</td>
<td>50-70% 1RM</td>
<td>NR</td>
<td>11 (8)</td>
<td>MRI</td>
</tr>
<tr>
<td>Johnson 2009</td>
<td>4 wks 3 times/wk 30-45 mins</td>
<td>A vs Control</td>
<td>50-70% peak VO₂</td>
<td>3,800</td>
<td>12 (7)</td>
<td>MRI</td>
</tr>
<tr>
<td>Levinger 2009</td>
<td>10 wks 3 times/wk 2-6 sets, 12-15 rep</td>
<td>R vs Control</td>
<td>75-85% 1RM</td>
<td>NR</td>
<td>HiMF: 15(15) LoMF: 12(13)</td>
<td>NA</td>
</tr>
<tr>
<td>Shojaee-Moradie 2007</td>
<td>6 wks 3 times/wk 20 mins</td>
<td>A vs Control</td>
<td>60-85% peak VO₂</td>
<td>3,586</td>
<td>10 (7)</td>
<td>MRS</td>
</tr>
<tr>
<td>Slentz 2011</td>
<td>34 wks 3 times/wk 19.2 km/wk</td>
<td>1. A vs R 2. A vs C</td>
<td>A: 75% peak VO₂ R: 70% 1RM</td>
<td>Not available</td>
<td>A: 48 R: 52 C: 44</td>
<td>CT</td>
</tr>
<tr>
<td>Sullivan 2012</td>
<td>16 wks</td>
<td>A vs Control</td>
<td>45-55% peak VO₂</td>
<td>21,325</td>
<td>12 (6)</td>
<td>MRS</td>
</tr>
<tr>
<td>Study</td>
<td>Duration</td>
<td>Frequency</td>
<td>Training Time</td>
<td>Intervention 1 vs Intervention 2</td>
<td>VO2 Contribution</td>
<td>VO2 Contribution</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Thompson 2010</td>
<td>24 wks</td>
<td>4 times/wk</td>
<td>30-60 mins</td>
<td>A vs Control</td>
<td>50-70% peak VO2</td>
<td>NR</td>
</tr>
<tr>
<td>Zelber-Sagi 2014</td>
<td>12 wks</td>
<td>3 times/wk</td>
<td>40 mins</td>
<td>R vs Control</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td><strong>Exercise and diet vs diet</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al-Jiffri 2013</td>
<td>12 wks</td>
<td>3 times/wk</td>
<td>30 mins</td>
<td>A vs Control</td>
<td>65-75% MHR</td>
<td>NR</td>
</tr>
<tr>
<td>Bozetto 2012</td>
<td>8 wks</td>
<td>2 times/wk</td>
<td>45 mins</td>
<td>Diet 1 &amp; A vs Diet 2</td>
<td>70% peak VO2</td>
<td>3,320</td>
</tr>
<tr>
<td>Goodpaster 2010</td>
<td>52 wks A or 26 wks Con</td>
<td>5 times/wk</td>
<td>60 mins</td>
<td>Diet &amp; A vs Diet</td>
<td>Diet &amp; A: 12 m moderate Diet: 6 m moderate</td>
<td>NR</td>
</tr>
<tr>
<td>Larson-Meyer 2008</td>
<td>26 wks</td>
<td>5 times/wk</td>
<td>individualized</td>
<td>Diet &amp; A vs Diet</td>
<td>NR</td>
<td>individualized</td>
</tr>
<tr>
<td>Promrat 2010</td>
<td>10,000 steps/day</td>
<td></td>
<td></td>
<td>Diet &amp; A vs Diet</td>
<td>Moderate</td>
<td>NA</td>
</tr>
<tr>
<td>Shah 2009</td>
<td>26 wks</td>
<td>3 times/wk</td>
<td>90 mins</td>
<td>Diet &amp; C vs Diet</td>
<td>70-85% MHR or 65-80% 1RM</td>
<td>23,400</td>
</tr>
<tr>
<td>Straznicky 2012</td>
<td>12 wks</td>
<td>3.5 times/wk</td>
<td>40 mins</td>
<td>Diet &amp; A vs Diet</td>
<td>65% MHR</td>
<td>15,294</td>
</tr>
<tr>
<td>Tamura 2005</td>
<td>2 wks</td>
<td>5-6 times/wk</td>
<td>2-3x30 min</td>
<td>Diet &amp; A vs Diet</td>
<td>50-60% peak VO2</td>
<td>NR</td>
</tr>
<tr>
<td>Yoshimura 2014</td>
<td>12 wks</td>
<td>3 times/wk</td>
<td>60 mins plus 120 mins/wk walking ≥300 min/wk</td>
<td>Diet &amp; A vs Diet</td>
<td>Moderate</td>
<td>3,067</td>
</tr>
</tbody>
</table>

**Legend:**

A – Aerobic exercise, BMI – Body Mass Index, C – Combined aerobic plus resistance exercise, Con – control, CT – Computed Tomography, HI:LO – high

Risk of Bias Assessment

The similarity in sample and effect sizes between randomized, controlled trials used in the analysis and those trials that were excluded from analysis suggests selection bias is minimal. The sample sizes are small however there is no apparent bias evident in the Egger plots (Supplementary Figures).
Discussion

This work analysed the effects of exercise training, with or without dietary intervention, on intrahepatic fat, body mass, body mass index, free fatty acids, insulin liver enzymes, lipids and cardiorespiratory fitness (peak VO₂). Our primary findings show that exercise training, in isolation from dietary intervention, has beneficial effects on some of these aforementioned outcome measures. We were able to establish that intrahepatic fat loss may be optimized by utilization of greater total exercise training energy expenditure. It appears that the addition of recently published work has added diversity to the current evidence base and increasingly varied study designs have necessitated a more cautious approach to data pooling and subsequent meta-analysis.

Our work failed to show changes in any of the three primary liver enzymes with exercise training. The number of participants included in these three analyses varied between 797 and 1109, so it unlikely that these analyses were underpowered. There were however several confounding variables in the included studies that may have prevented the isolated effects of exercise to be examined. There were eight studies that, in addition, to exercise training also provided various dietary interventions. Moreover, the exercise training programs varied greatly between studies with respect to exercise intensity, duration, frequency and modality (eg aerobic versus resistance training). Whatever the reason for the absence of significant changes, our work confirms the previous findings that liver enzymes have limited sensitivity in detecting early liver disease[9]. It appears that liver enzymes are, at best, a blunt tool for assessing change in liver function after exercise training.

Our analysis suggested that intrahepatic fat is reduced with exercise training, in as little as 4 weeks[27], but aerobic exercise programs of total energy expenditure >10,000Kcal are likely to elicit greater improvements than programs with lower exercise energy expenditure. Our findings in this respect advance those of Keating et al.[6] who conducted pooled analyses of all exercise modalities.
In mild fatty liver disease, the first choice therapy is lifestyle modification, although medications that decrease insulin resistance, hyperlipidaemia, and those that induce weight loss have been shown to improve liver function [53]. Our analysis reinforces the fact that even short term, exercise training is effective in reducing intrahepatic fat. These findings suggest that exercise may benefit liver health through a calorie expenditure mechanism, sufficient levels of which are acquired more rapidly with vigorous or high intensity exercise training. Data from heart failure studies suggests that high intensity exercise may promote superior health benefits [11].

It has been recognized that free fatty acids are the vehicle by which triacylglycerol is stored in adipose tissue and is transported to its site of utilization [54]. Free fatty acid turnover is rapid with a half-life of 2-4 minutes [54]. In our analysis, 3 studies using >10,000Kcal total exercise program energy expenditure appear to elicit greater reductions in intrahepatic fat. High volume exercise energy expenditure provides a variety of benefits related to metabolic disorders; we propose that one clinical benefit directly related to liver function is that high volume aerobic exercise elicits triacylglycerol consummation. Of course other exercise-induced benefits will accrue and primarily affect other organs, besides the liver, such as improved flow mediated vascular dilatation [55], cardiac function [56], peak VO$_2$ [55, 56] and weight loss which was minimal in our analysis.

Our analyses suggest that it is calorie burning that elicits reductions in liver fat, reduced fat storage, possibly reduced liver enzymes and the sum of all of this is improved glycaemic control. Of perhaps most interest is that the above changes are achieved almost totally independent of body mass changes. To illustrate this we chose the study of Shoajee-Moradie [32], because this work was the only included study to utilize high intensity exercise and measure peak VO$_2$, allowing calorie expenditure to be calculated. The exercise participants in Shoajee-Moradie’s study [32] expended 9.8 Kcal per minute, or 590 Kcal each week, the study only ran for 6 weeks. Exercise participants therefore only expended 3500 Kcal during the whole study, equivalent to 0.5 kg of fat. Not surprisingly, Shoajee-Moradie reported no change in body mass in either
exercise or control groups[32]. Nevertheless Johnson’s study[27] reported a similar aggregate exercise program energy expenditure (3800Kcal) but as it was delivered in a shorter 4-week period, changes in interhepatic fat were noted.

**Limitations**

A major limitation of this work was that considerable heterogeneity meant that data pooling was unjustified in a number of meta-analyses. We systematically attempted to identify reasons for heterogeneity by grouping studies according to similarities in interventions, exercise programs, patient populations, but we failed to identify definitive conclusions. We were able to reduce heterogeneity somewhat by limiting data pooling to studies that did not use concurrent dietary interventions. Some patients were overweight/obese while others had definitive liver disease, some of the former may well have also exhibited clinical NAFLD. The likelihood is that the potential for non-clinical NAFLD participants to ‘regress to the mean’ was probably less than in those with diagnosed NAFLD.

As discussed previously liver enzymes have limited predictive value for liver function and it is notable that only one of the included studies used the gold standard liver biopsy method to assess liver function. There were studies that, in addition, to exercise training also provided various dietary interventions. The exercise training programs varied greatly between studies with respect to exercise intensity, duration, frequency and modality (eg aerobic versus resistance training). The normal distribution of the Egger plots evidenced minimal risk of publication bias.

In the case of the liver enzymes, some heterogeneity may be explained by the naturally occurring inter-subject variation that exists with these enzymes. Measures of lean, and fat, body mass would have shed more light onto the role that body composition plays in improving liver function through exercise. We would like to have conducted sub-group analyses of those with normal and impaired liver function but group-level (as opposed to individual patient data) precluded this. We considered program duration, modality and frequency sub-analyses but the variation in included study data precluded this.
Meta-regression between change in liver function and body mass was considered, but change in body mass was not reported in all studies.

The current number of published studies does not allow meta-analyses to compare; aerobic versus resistance exercise; high intensity versus low or moderate intensity exercise; and effects of weekly exercise duration. All of which may influence change in current guidelines.

The risk of bias assessment identified that allocation concealment to investigators was conducted in less than one quarter if studies and investigator bias is quite likely. Moreover assessor blinding was uncommon (only 5 of 21 studies), so this would not mitigate allocation concealment. The lack of physical activity monitoring in 50% of the studies implies that there was likelihood that sedentary controls crossed over to exercise in some studies. The lack of assessment of energy expended during exercise in 10 studies meant that our meta-analyses of aggregate energy expenditure may not be fully representative.

Future studies should aim to at least blind the outcome assessors to the participants’ intervention to minimise outcome assessment bias. Future work may also wish to focus on exercise induced calorie expenditure and the presence or absence of associated weight loss, in order to clarify the mechanism of benefit. Use of gold standard biopsy, rather than reliance on liver enzymes, assessment of liver function is recommended.

Conclusions

Exercise training reduces intrahepatic fat and free fatty acids while increasing cardiorespiratory fitness. An aggregate exercise program energy expenditure (>10,000 Kcal) may be required to promote reductions in intrahepatic fat.

Acknowledgements

None
References


**Figures and Tables**

Figure 1. PRISMA Statement

Figure 2. Study Quality Assessment

Key: Total out of 15 Points

Legend: # Three points possible – 1 point if adherence>85%, 1 point if adverse events reported, 1 point if exercise attendance is reported

*Two points possible – 1 point if primary outcome is reported, 1 point if all other outcomes reported
Figure 3 Percentage change in intrahepatic fat: Exercise Programs >10,000 Kcal energy expenditure.

Figure 4. Change in Body Mass Index (BMI): Exercise Interventions

Figure 5. Change in Fasting free fatty acids (FFA): Exercise Interventions

Figure 6. Change in Gamma-Glutamyl Transpeptidase (GGT): Exercise Interventions

Figure 7. Change in Insulin: Exercise Interventions

Figure 8. Change in Total Cholesterol: Exercise Interventions

Table 1. Table of Outcome Measures

Table 2. Included Studies – Exercise Programming Details

Supplementary Files

Figure S1. PubMed Search Strategy

Egger Plots to Examine Publication Bias

Figure S2. Change in Interhepatic Fat in Exercise Studies with >10,000 Kcal Total Energy Expenditure.

Figure S3. Change BMI in Exercise Studies Only

Figure S4. Change in FFA in Exercise Studies Only

Figure S5. Change in GGT in Exercise Studies Only

Figure S6 Change in Insulin in Exercise Studies Only

Figure S7. Change in Total Cholesterol in Exercise Studies Only

Table S1. Excluded Randomized Controlled Trials