The effect of exercise on quality of life and activities of daily life in frail older adults: a systematic review of randomised control trials.

Evan Campbell*[a], Fanny Petermann-Rocha*[b,c], Paul Welsh[c], Carlos Celis-Morales[b,c], Jill P Pell[b], Frederick K Ho[b], and Stuart R Gray[c].

[a]Healthcare Improvement Scotland, NHSScotland

[b]Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK.

[c]British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Science, University of Glasgow, Glasgow, UK

* EC and FP-R contributed equally to this work and are joint first authors.

Corresponding author

Dr Evan Campbell

Delta House, 50 West Nile Street, Glasgow G1 2NP

evan.campbell@nhs.scot

Abstract (100 to 200 words)

While the positive effects of exercise on frailty are well documented, the effect of exercise on quality of life (QoL) and activities of daily living (ADL) in frail older adults remains less certain. Therefore, this paper aimed to systematically review the literature investigating the effect of exercise on QoL and ADL in this group. Embase, MEDLINE, CENTRAL, PEDro and Web of Science Core Collections were searched systematically using relevant MeSH terms. The inclusion criteria were: controlled trial design, published in English, population included frail older adults, frailty measured quantitatively, interventions that included exercise, and QoL or ADL measurements (PROSPERO: CRD42018106173). After screening, 15 studies were eligible for inclusion in the qualitative synthesis (total n: 2,467; mean age range: 70 - 85 years). There was a positive effect on QoL or ADL measures in 10 out of the 15 studies. QoL and ADLs only improved in studies that also reported improved physical outcomes. These results reflect the multi-factoral nature of frailty and how physical capability and QoL are interlinked. Heterogeneity precluded formal meta-analysis. Future trials in frail older adults should focus on interventions that include exercise, measure physical outcomes and use consistent study design to enable meta-analysis to be conducted.

Keywords: Frailty; Ageing; Quality of life; Activities of daily living.

Highlights

- Exercise interventions for frail older adults rarely focus on QoL and ADLs
- There was no obvious link between type of exercise and improvements in QoL and ADLs
- Improvements in QoL and ADLs were only seen with improvements in physical outcomes
- Future clinical trials are still needed to conduct a proper meta-analysis.

1. Introduction

Frailty is a clinical state which primarily affects older adults, places the individual at a high risk of hospital admission and warrants clinical intervention (Fried et al., 2001). Frailty also has negative outcomes on hospitalisation, independence and thus can place a burden on public health care systems. In the United Kingdom, approximately 6.5% of community dwelling older adults aged between 60-69 years are frail, and the prevalence rises tenfold with age to approximately 65% of older adults aged 90 years or more being frail (Gale et al., 2015). Furthermore, the proportion of those living with severe or moderate frailty in the UK is 3 and 12%, respectively (BMA, 2020). The clinical phenotype of frailty comprises both psychological and physical components evaluating mood, physical condition and physical function (Fried et al., 2001). Frailty has a negative impact on activities of daily living (ADL) and instrumental ADL (IADL) with approximately 60% of those with frailty being negatively impacted compared to approximately 14% of non-frail older adults (Gale et al., 2015). A systematic review and meta-analysis demonstrated an inverse association between quality of life (QoL) and frailty among community-dwelling older people (Kojima et al., 2016). The same review proposed that health interventions aimed at reducing frailty may also have a positive impact on QoL (Kojima et al., 2016). Therefore, ADL and QoL have an important impact in older adults and in geriatric medicine.

Fried et al. defined frailty as a clinical phenotype in 2001 and at the time noted the negative impact that frailty has on QoL (Fried et al., 2001). Since its inception as a clinical phenotype, there have been numerous operational classification tools for frailty (Bouillon et al., 2013). However, it has been shown that not all measurements of frailty are appropriate for all purposes (Martin and Brighton, 2008). A common theme among definitions is a quantitative decrease in physical function or independence (Dent et al., 2016; Devereux et al., 2019), making exercise a potentially effective therapy.

Supervised exercise is safe in frail older adults (de Labra et al., 2015) and is recommended as a possible therapy to reverse frailty and increase independence (Aguirre and Villareal, 2015). A previous systematic review, conducted by Chou et al., assessed the effect of exercise training on ADL and QoL in frail older adults (Chou et al., 2012). The main finding of that systematic review was that there was no effect of exercise on QoL with a small positive effect on ADL observed; however, the results were far from conclusive due to the poor quality and low volume of the evidence (Chou et al., 2012). In addition, the literature search of that systematic review was undertaken in 2010, and since then, there has been a considerable amount of primary research concerning exercise and frailty published. Therefore, another review of the literature is warranted. The current paper aims to

systematically review the literature investigating the effect of exercise on QoL and ADL in older adults who were clinically frail. Furthermore, we wanted to explore if different types of exercise had different effects on QoL in frail older adults and if there were any patterns in the evidence between changes in physical outcomes and changes in ADLs or QoL.

2. Methods

This systematic review was conducted according to the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). The protocol for this systematic review was published in PROSPERO (ID protocol: CRD42018106173).

2.1. Data sources and search strategy

A systematic search was undertaken in August 2018 using the following databases: MEDLINE, EMBASE, CENTRAL and Web of Science Core collections. Relevant MeSH terms and Boolean phases were used for the search: frail$ AND exercis$ OR train$ OR strength$ without time restriction. QoL and ADL were not included in the initial search strategy as these are often measured as secondary outcomes and therefore, may be not present in abstracts or titles. The complete search strategy is shown in the appendix.

2.2. Study selection

Randomised controlled trials (RCT) or non-randomised clinical trials conducted in frail older people (defined clinically or quantitatively, e.g. using the Fried phenotype (Fried et al., 2001), frailty index or a defined decrement in an ADL measure) with a control group and published in English were considered for eligibility. Study populations could be community dwelling or nursing home residents. Non-human studies, studies including participants younger than 65-years-old or being described as frail with no quantitative measurement of frailty, e.g. all hip replacement patients, and studies without a control group were excluded.

2.3. Data extraction

All articles were screened for suitability by title and abstract by two reviewers (EC and FP-R). If there was insufficient information to gauge eligibility, the methods were screened. Full articles were screened by the same researchers (EC and FP-R). If information was missing or there was ambiguity, the lead author was contacted for clarification. Data from eligible articles were extracted using a standardised form. The form included: authors, date of publication, study design, sample size (intervention and control group), age of participants, was the sample satistically powered, was analysis carried out on the basis of intention to treat, attrition rate, length of intervention, length of

follow-up, frailty measure, type of intervention, frequency and type of exercise, ADL/QoL measure, other relevant outcomes measures and results.

2.4. Methodological quality assessment

Studies included in the systematic review were further assessed using the PEDro scale. This instrument evaluates both internal and external validity, gives a score out of 10 and is a reliable scale of the methodological quality of randomised control trials (de Morton, 2009). Each article was scored independently by two reviewers (EC and FP-R), and divergent scores were settled by discussion. Quality assessment was based on the content of the study in the published article. Due to the nature of frailty exercise studies often using non-frail participants as controls, this removed the possibility of randomisation from these studies. Due to this, there was no lower limit placed on PEDro scores.

2.5. Research questions

The primary research question was:

Does exercise have an effect on QoL or ADL in frail older adults?

The secondary research question was:

Does the type of exercise intervention, or the measure of frailty, impact the effect of exercise on QoL or ADL measures in frail older adults?

3. Results

3.1. Search results

From the search, 6,395 records were identified. After the removal of 152 duplicates, 6,243 titles and abstracts were screened, 5,695 of which were excluded (Fig. 1). Full texts of 548 studies were read for further information, and further 531 were excluded (Fig. 1). Fifteen studies, described by 17 articles, were included in the qualitative synthesis. In the evidence synthesis, only the primary articles are cited unless describing results specifically from the secondary article. Due to the variation in outcome measures and interventions prescribed, a meta-analysis was not appropriate.

INSERT FIGURE ONE HERE (PRISMA DIAGRAM).

3.2. Quality of studies

PEDro scale scores ranged from three (Rydwik et al., 2010) to eight points (El-Khoury et al., 2015; Fairhall et al., 2012; Helbostad et al., 2004; Hoogeboom et al., 2010; Latham et al., 2003; Littbrand et al., 2009; Talley et al., 2017; Villareal et al., 2006) (Table 1). Scores were most commonly low due to lack of blinding of the participants and therapists delivering the intervention, which is unavoidable in exercise trials (Campbell et al., 2016; Campbell et al., 2018). Studies were not excluded due to a low PEDro score.

INSERT TABLE 1 HERE (PEDRO SCORES).

3.3. Study characteristics

All studies included were RCTs, no eligible non randomised controlled trials were indentified. The total number of participants from the studies included in the review was 2,467, and the mean age ranged from 70 to 85 years (Table 2) (Cameron et al., 2013; El-Khoury et al., 2015; Giné-Garriga et al., 2010; Grönstedt et al., 2013; Helbostad et al., 2004; Hoogeboom et al., 2010; Langlois et al., 2013; Latham et al., 2003; Littbrand et al., 2009; Rydwik et al., 2010; Talley et al., 2017; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2006; Villareal et al., 2011). Sample size ranged from 21 (Hoogeboom et al., 2010) to 706 participants (El-Khoury et al., 2015). Nine studies provided a sample size calculation (Cameron et al., 2013; El-Khoury et al., 2015; Grönstedt et al., 2013; Helbostad et al., 2004; Latham et al., 2003; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2006; Villareal et al., 2011) but only two of these calculations were powered for a QoL or ADL outcome (Grönstedt et al., 2013; Latham et al., 2003). Six studies used a clinically validated measure of frailty (Cameron et al., 2013; Hoogeboom et al., 2010; Langlois et al., 2013; Latham et al., 2003; Rydwik et al., 2010; Tarazona-Santabalbina et al., 2016).

Eight studies included an intervention that was solely exercise (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Helbostad et al., 2004; Hoogeboom et al., 2010; Langlois et al., 2013; Latham et al., 2003; Littbrand et al., 2009; Tarazona-Santabalbina et al., 2016), five studies had a multi-modal intervention which included an exercise component (Cameron et al., 2013; Grönstedt et al., 2013; Talley et al., 2017; Villareal et al., 2017; Villareal et al., 2006), while two studies were multi-armed RCTs which included an interventional arm which was exercise only and another which was multi-modal (Rydwik et al., 2010; Villareal et al., 2011). Other components in multi-modal interventions included a dietary or nutrition intervention (n=5); physical, cognitive or social goal-setting (n=2); physiotherapy (n=1); geriatrician input (n=1); rehabilitation physician input (n=1); and bladder rehabilitation (n=1) (Table 1).

All exercise interventions, bar one (Latham et al., 2003), were multi-component exercise interventions. All studies included strength training in their intervention, 11 included balance training (Cameron et al., 2013; El-Khoury et al., 2015; Giné-Garriga et al., 2010; Grönstedt et al., 2013; Helbostad et al., 2004; Langlois et al., 2013; Littbrand et al., 2009; Rydwik et al., 2010; Tarazona-Santabalbina et al., 2016; Villareal et al., 2006; Villareal et al., 2011), nine included aerobic training (Cameron et al., 2013; Grönstedt et al., 2013; Hoogeboom et al., 2010; Langlois et al., 2013; Rydwik et al., 2010; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2006; Villareal et al., 2011), four included flexibility exercises (El-Khoury et al., 2015; Tarazona-Santabalbina et al., 2016; Villareal et al., 2006; Villareal et al., 2011), four included functional exercises (Cameron et al., 2013; Hoogeboom et al., 2010; Littbrand et al., 2009; Rydwik et al., 2010), and two included walking (Grönstedt et al., 2013; Littbrand et al., 2009). Thirteen of the exercise interventions were supervised (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Grönstedt et al., 2013; Helbostad et al., 2004; Hoogeboom et al., 2010; Langlois et al., 2013; Littbrand et al., 2009; Rydwik et al., 2010; Talley et al., 2017; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2006; Villareal et al., 2011) while two were unsupervised home exercise plans (Cameron et al., 2013; Latham et al., 2003).

The lowest frequency of training was twice a week (Giné-Garriga et al., 2010; Helbostad et al., 2004; Hoogeboom et al., 2010; Rydwik et al., 2010) and the highest five times a week (Cameron et al., 2013; Talley et al., 2017; Tarazona-Santabalbina et al., 2016). Nine studies reported intensity of the exercise prescribed which was typically between 60-80% of maximal capability (Giné-Garriga et al., 2010; Helbostad et al., 2004; Hoogeboom et al., 2010; Latham et al., 2003; Littbrand et al., 2009; Rydwik et al., 2010; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2011). Seven of those progressed the intensity of exercises to keep them challenging (Giné-Garriga et al.,

2010; Hoogeboom et al., 2010; Latham et al., 2003; Littbrand et al., 2009; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2011), one did not progress the exercise (Talley et al., 2017) and one did not report if there was progression (Rydwik et al., 2010). Six studies did not report the intensity of their exercise intervention (Cameron et al., 2013; El-Khoury et al., 2015; Grönstedt et al., 2013; Helbostad et al., 2004; Langlois et al., 2013; Villareal et al., 2006), but two of these did report that they progressed exercises so that they remained challenging (Cameron et al., 2013; Grönstedt et al., 2013). Trial intervention length ranged from three weeks (Hoogeboom et al., 2010) to two years (El-Khoury et al., 2015); however, the most common length of the intervention was 12 weeks which was employed in seven studies (Giné-Garriga et al., 2010; Grönstedt et al., 2013; Helbostad et al., 2004; Langlois et al., 2013; Littbrand et al., 2009; Rydwik et al., 2010; Talley et al., 2017).

Thirteen different QoL and ADL outcome measures and 37 other outcome measures of varying domains were used across the 15 studies (Table 2).

3.4. Primary research question: Does exercise have an effect on QoL or ADL in frail older adults?

The superiority of the exercise intervention over control for QoL or ADL, was reported in nine studies (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Helbostad et al., 2004; Langlois et al., 2013; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2011) (Grönstedt et al., 2013; Villareal et al., 2006) with varying effect sizes for QoL and ADL (Table A-1). From these nine studies, between-group improvements in the exercise group, or maintenance in the exercise group and with deterioration in the control group, in QoL or ADL measures, were reported in seven studies (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Helbostad et al., 2004; Langlois et al., 2013; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2011). In two studies, there were within-group improvements in the exercise group (Grönstedt et al., 2013; Villareal et al., 2006). In one study, there were improvements seen in both the intervention and control groups (Talley et al., 2017) with no overall superiority in the exercise group. However, in five studies, there was no effect on QoL or ADL in either the exercise group or control groups (Cameron et al., 2013; Hoogeboom et al., 2010; Latham et al., 2003; Littbrand et al., 2009; Rydwik et al., 2010). In summary, there was a positive effect on QoL or ADL measures in 10 out of the 15 studies included in the review. It is, therefore, unclear if exercise had an effect on QoL or ADL in frail older adults.

3.5. Secondary research question: Does the type of exercise intervention, or the measure of frailty, impact the effect of exercise on QoL or ADL measures in frail older adults?

Of the nine studies which observed improvement or maintenance from the exercise intervention on QoL or ADL measures, four used a clinically validated measure of frailty (Langlois et al., 2013; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2006) and five did not (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Grönstedt et al., 2013; Helbostad et al., 2004; Villareal et al., 2011) (Table 2). Similarly, in the six studies which did not observe an improvement, four employed a clinically recognised frailty measure (Cameron et al., 2013; Hoogeboom et al., 2010; Latham et al., 2003; Rydwik et al., 2010) while two did not (Littbrand et al., 2009; Talley et al., 2017). From the studies included in this review, there was no evidence to suggests that using a clinically recognised measure of frailty impacted on the effect of exercise on QoL or ADLs.

All exercise interventions bar one (Latham et al., 2003), were multi-component exercise interventions and all studies included strength training. Due to the large variation in exercise interventions implemented (Table A.1.) and the heterogeneity of the results, it is unclear if the type of exercise used would have different effects on QoL and ADL measures in frail older adults.

Of the nine studies where improvements were reported, six utilised an exercise-only intervention (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Helbostad et al., 2004; Langlois et al., 2013; Tarazona-Santabalbina et al., 2016; Villareal et al., 2011) and three were part of a multi-modal intervention (Grönstedt et al., 2013; Villareal et al., 2017; Villareal et al., 2006). The other components in the multi-modal interventions included a dietary intervention (Villareal et al., 2017; Villareal et al., 2006) and goal setting (Grönstedt et al., 2013). Similarly, of the trials in which no positive effects of exercise were observed, four were exercise-only interventions (Hoogeboom et al., 2010; Latham et al., 2003; Littbrand et al., 2009; Rydwik et al., 2010) and two were part of multi-modal interventions (Cameron et al., 2013; Talley et al., 2017). The other components of the multi-modal interventions were bladder rehabilitation (Talley et al., 2017) and a multi-disciplinary intervention including input from a physiotherapist, dietician, geriatrician, a rehabilitation specialist, and goal setting (Cameron et al., 2013) . In summary, there was no evidence to suggest that the effectiveness of exercise depended on whether it was prescribed as part of a multi-modal intervention or by itself on QoL or ADL measures in frail older adults.

All nine studies which observed a positive effect of exercise on QoL or ADL also observed an improvement in physical outcome measures (El-Khoury et al., 2015; Giné-Garriga et al., 2010; Grönstedt et al., 2013; Helbostad et al., 2004; Langlois et al., 2013; Tarazona-Santabalbina et al., 2016; Villareal et al., 2017; Villareal et al., 2006; Villareal et al., 2011). Of the six studies that did not

250 observe a positive effect of exercise on ADL or QoL measures, four reported an improvement in
251 physical outcome measures (Cameron et al., 2013; Littbrand et al., 2009; Rydwik et al., 2010; Talley
252 et al., 2017) and two did not (Hoogeboom et al., 2010; Latham et al., 2003). From the studies
253 included in the review, it would appear that improvements in physical function do not necessarily
254 result in improvements in ADL and Qol, but when these are increased, physical function is also
255 increased.

256

257 ## 4. Discussion

258 ### *4.1. Main findings*

259 This systematic review identified 15 RCTs that examined the effect of exercise on QoL or ADLs in frail
260 older adults. Overall, the published studies produced inconsistent findings on whether prescribed
261 exercise led to a positive effect on QoL or ADLs in frail older adults. This lack of consistency is likely
262 to reflect heterogeneity in study design and the interventions evaluated. Interestingly QoL and ADLs
263 only improved in studies that also reported improved physical outcomes. There was no evidence
264 that using a clinically validated measure of frailty impacted the results.

265 ### *4.2. Comparison with previous studies*

266 The findings of this review are broadly in line with the previous systematic review and meta-analysis
267 which also concluded that there was no noticeable effect of exercise on QoL (Chou et al., 2012).
268 Chou et al. did observe, in their meta-analysis, an improvement in ADLs after exercise training (Chou
269 et al., 2012); however, this result was based on three trials with heterogeneity in the intervention
270 prescribed and outcomes measured. Likewise, in this present review, large heterogeneity was
271 observed in the interventions prescribed and the outcome measures used. For these reasons, a
272 meta-analysis of the results would have been inappropriate (Higgins et al., 2019). The authors of a
273 systematic review, published in 2011, examining the effect of exercise on the management of frailty
274 also chose not to conduct a meta-analysis of their results when examining the effect of exercise on
275 QoL or ADLs in frail older adults (Theou et al., 2011). These authors also concluded that it was not
276 possible from the included literature to determine an effect or not. Subsequently, it is recommended
277 that future trials for frail older adults should be more consistent in regard to intervention
278 prescription and outcome measures used. Such homogeneity with a primary focus on ADLs and QoL
279 would allow a robust meta-analysis to draw a firm conclusion.

280 ### *4.3. Implications of the findings*

281 Frailty is theoretically defined as a clinically recognisable state of increased vulnerability resulting
282 from an ageing-associated decline in reserve and function across multiple physiologic systems, such
283 that the ability to cope with everyday or acute stressors is comprised. Since Fried defined the clinical
284 phenotype in 2001 (Fried et al., 2001), there has been a plethora of frailty measures and indices
285 created (Walston et al., 2018). Indeed, in the review by Chou et al., the authors created their own
286 definition of frailty based on the Fried criteria and used this definition as an inclusion criterion (Chou
287 et al., 2012). Given the large number of frailty measures available, research is needed to compare
288 these recognised measures and 'ad hoc' quantitative measures to investigate if using a recognised

289 measure is in fact necessary or if simply measuring a decrease in physical and psychological function
290 is sufficiently sensitive to identify frail older adults.

291 To our knowledge, this is the first review to date to question whether using a recognised measure of
292 frailty impacts upon findings around the effect of exercise. There was no evidence of a difference in
293 whether there was a positive effect, or not, of exercise on QoL or ADLs in studies which included a
294 clinically recognised measure or frailty or not. Additionally, there was some evidence that
295 improvements in QoL and ADLs may be linked to improvements in physical outcomes. In this
296 systematic review, it was observed that there was a positive effect on QoL or ADLs in all of the
297 studies that also reported an improvement in a physical outcome measure (e.g. balance) and none
298 of the studies that did not. This is logical if we consider that frailty is, by definition, multifactorial
299 (Fried et al., 2001); therefore, addressing the physical aspects of frailty will have an impact on QoL
300 and ADLs aspect. This may indicate that the improvements in physical function do not necessarily
301 result in improvements in ADL and Qol, but when these are increased, physical function is increased
302 as well. This is a novel finding when it is compared to previous reviews (Chou et al., 2012; Theou et
303 al., 2011) and with an umbrella review of systematic reviews which found that exercise seemed to
304 have a positive effect on QoL frail older adults; however, the optimum exercise program remained
305 still unclear (de Labra et al., 2015). Consequently, this theory could be tested robustly if there is
306 greater homogeneity in interventions, outcome measures and an optimum exercise programme is
307 identified. A hypothetical pathway of the influence of exercise on physical function, frailty and QoL
308 and ADLs can be seen in figure A.1.

309 *4.4. Strengths and limitations*

310 This work was conducted in accordance with the recommendations from the PRISMA statement
311 (Moher et al., 2009). Study selection, data extraction and assessment of quality were carried out by
312 two independent reviewers (Higgins et al., 2019). However, one of the main limitations of this study
313 was the inclusion of articles in English only. This could have limited the involvement of other frailty
314 populations from other regions (e.g., China and Latin America) and generated a language bias. In
315 addition, the heterogeneity of the studies in terms of intervention and study methodology
316 prohibited the derivation of pooled estimates from a meta-analysis. Such limitations are
317 commonplace in interventions which have such large possible variations such as exercise regimes. As
318 with all qualitative syntheses, any comparisons made are observational, and causality cannot be
319 inferred. Finally, the literature search was carried out two years ago and may not reflect the most
320 recent literature. However, the last set of authors to review this topic carried out their literature

search more than 10 years ago, therefore this present document still makes a valid contribution to the literature.

4.5. Conclusions

In conclusion, heterogeneity in study design and the intervention evaluated has produced inconsistent findings, to date, as to whether prescribed exercise improves QoL or ADLs in the frail elderly. However, further scrutiny suggests that, to be effective in QoL or ADLs, exercise interventions may have generated a positive effect on physical outcomes too. Future trials should use a consistent methodology and focus on this type of intervention so that robust conclusions can be drawn.

Declaration of interest

None

Author contributions

EC performed the search strategy and study selection. EC and F.P-R screened all articles, data extraction and the quality assessment. EC and F.P-R wrote the first version of the manuscript. All authors critically reviewed this and previous drafts. All authors approved the final draft for submission. EC is the guarantor.

Funding

FP-R receives financial support from the Chilean Government for doing her PhD (ANID-Becas Chile 2018 – 72190067).

References

Aguirre, L.E., Villareal, D.T., 2015. Physical Exercise as Therapy for Frailty. Nestle Nutr Inst Workshop Ser 83, 83-92.

BMA, 2020. Identification and management of patients with frailty.

Bouillon, K., Kivimaki, M., Hamer, M., Sabia, S., Fransson, E.I., Singh-Manoux, A., Gale, C.R., Batty, G.D., 2013. Measures of frailty in population-based studies: an overview. BMC Geriatr 13, 64.

Cameron, I.D., Fairhall, N., Langron, C., Lockwood, K., Monaghan, N., Aggar, C., Sherrington, C., Lord, S.R., Kurrle, S.E., 2013. A multifactorial interdisciplinary intervention reduces frailty in older people: randomized trial. BMC Medicine 11, 65.

Campbell, E., Coulter, E.H., Mattison, P.G., Miller, L., McFadyen, A., Paul, L., 2016. Physiotherapy Rehabilitation for People With Progressive Multiple Sclerosis: A Systematic Review. Archives of Physical Medicine and Rehabilitation 97, 141-151.e143.

Campbell, E., Coulter, E.H., Paul, L., 2018. High intensity interval training for people with multiple sclerosis: A systematic review. Mult Scler Relat Disord 24, 55-63.

Chou, C.-H., Hwang, C.-L., Wu, Y.-T., 2012. Effect of Exercise on Physical Function, Daily Living Activities, and Quality of Life in the Frail Older Adults: A Meta-Analysis. Archives of Physical Medicine and Rehabilitation 93, 237-244.

de Labra, C., Guimaraes-Pinheiro, C., Maseda, A., Lorenzo, T., Millán-Calenti, J.C., 2015. Effects of physical exercise interventions in frail older adults: a systematic review of randomized controlled trials. BMC Geriatrics 15, 154.

de Morton, N.A., 2009. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. Aust J Physiother 55, 129-133.

Dent, E., Kowal, P., Hoogendijk, E.O., 2016. Frailty measurement in research and clinical practice: A review. Eur J Intern Med 31, 3-10.

Devereux, N., Ellis, G., Dobie, L., Baughan, P., Monaghan, T., 2019. Testing a proactive approach to frailty identification: the electronic frailty index. BMJ Open Qual 8, e000682-e000682.

El-Khoury, F., Cassou, B., Latouche, A., Aegerter, P., Charles, M.-A., Dargent-Molina, P., 2015. Effectiveness of two year balance training programme on prevention of fall induced injuries in at risk women aged 75-85 living in community: Ossébo randomised controlled trial. BMJ : British Medical Journal 351, h3830.

Fairhall, N., Sherrington, C., Kurrle, S.E., Lord, S.R., Lockwood, K., Cameron, I.D., 2012. Effect of a multifactorial interdisciplinary intervention on mobility-related disability in frail older people: randomised controlled trial. BMC Medicine 10, 120.

Fried, L.P., Tangen, C.M., Walston, J., Newman, A.B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W.J., Burke, G., McBurnie, M.A., 2001. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 56, M146-156.

Gale, C.R., Cooper, C., Sayer, A.A., 2015. Prevalence of frailty and disability: findings from the English Longitudinal Study of Ageing. Age Ageing 44, 162-165.

Giné-Garriga, M., Guerra, M., Pagès, E., Manini, T.M., Jiménez, R., Unnithan, V.B., 2010. The effect of functional circuit training on physical frailty in frail older adults: a randomized controlled trial. J Aging Phys Act 18, 401-424.

Grönstedt, H., Frändin, K., Bergland, A., Helbostad, J.L., Granbo, R., Puggaard, L., Andresen, M., Hellström, K., 2013. Effects of individually tailored physical and daily activities in nursing home residents on activities of daily living, physical performance and physical activity level: a randomized controlled trial. Gerontology 59, 220-229.

Helbostad, J.L., Sletvold, O., Moe-Nilssen, R., 2004. Home training with and without additional group training in physically frail old people living at home: effect on health-related quality of life and ambulation. Clin Rehabil 18, 498-508.

Higgins, J.P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A., 2019. Cochrane handbook for systematic reviews of interventions. John Wiley & Sons.

Hoogeboom, T.J., Dronkers, J.J., van den Ende, C.H., Oosting, E., van Meeteren, N.L., 2010. Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized pilot trial. Clin Rehabil 24, 901-910.

Kojima, G., Iliffe, S., Jivraj, S., Walters, K., 2016. Association between frailty and quality of life among community-dwelling older people: a systematic review and meta-analysis. J Epidemiol Community Health 70, 716-721.

Langlois, F., Vu, T.T., Chassé, K., Dupuis, G., Kergoat, M.J., Bherer, L., 2013. Benefits of physical exercise training on cognition and quality of life in frail older adults. J Gerontol B Psychol Sci Soc Sci 68, 400-404.

Latham, N.K., Anderson, C.S., Lee, A., Bennett, D.A., Moseley, A., Cameron, I.D., 2003. A randomized, controlled trial of quadriceps resistance exercise and vitamin D in frail older people: the Frailty Interventions Trial in Elderly Subjects (FITNESS). J Am Geriatr Soc 51, 291-299.

Littbrand, H., Lundin-Olsson, L., Gustafson, Y., Rosendahl, E., 2009. The effect of a high-intensity functional exercise program on activities of daily living: a randomized controlled trial in residential care facilities. J Am Geriatr Soc 57, 1741-1749.

Martin, F.C., Brighton, P., 2008. Frailty: different tools for different purposes? Age Ageing 37, 129-131.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339, b2535.

Napoli, N., Shah, K., Waters, D.L., Sinacore, D.R., Qualls, C., Villareal, D.T., 2014. Effect of weight loss, exercise, or both on cognition and quality of life in obese older adults. Am J Clin Nutr 100, 189-198.

Rydwik, E., Frändin, K., Akner, G., 2010. Effects of a physical training and nutritional intervention program in frail elderly people regarding habitual physical activity level and activities of daily living--a randomized controlled pilot study. Arch Gerontol Geriatr 51, 283-289.

Talley, K.M.C., Wyman, J.F., Bronas, U., Olson-Kellogg, B.J., McCarthy, T.C., 2017. Defeating Urinary Incontinence with Exercise Training: Results of a Pilot Study in Frail Older Women. J Am Geriatr Soc 65, 1321-1327.

Tarazona-Santabalbina, F.J., Gómez-Cabrera, M.C., Pérez-Ros, P., Martínez-Arnau, F.M., Cabo, H., Tsaparas, K., Salvador-Pascual, A., Rodriguez-Mañas, L., Viña, J., 2016. A Multicomponent Exercise Intervention that Reverses Frailty and Improves Cognition, Emotion, and Social Networking in the Community-Dwelling Frail Elderly: A Randomized Clinical Trial. J Am Med Dir Assoc 17, 426-433.

Theou, O., Stathokostas, L., Roland, K.P., Jakobi, J.M., Patterson, C., Vandervoort, A.A., Jones, G.R., 2011. The effectiveness of exercise interventions for the management of frailty: a systematic review. J Aging Res 2011, 569194.

Villareal, D.T., Aguirre, L., Gurney, A.B., Waters, D.L., Sinacore, D.R., Colombo, E., Armamento-Villareal, R., Qualls, C., 2017. Aerobic or Resistance Exercise, or Both, in Dieting Obese Older Adults. New England Journal of Medicine 376, 1943-1955.

Villareal, D.T., Banks, M., Sinacore, D.R., Siener, C., Klein, S., 2006. Effect of weight loss and exercise on frailty in obese older adults. Arch Intern Med 166, 860-866.

Villareal, D.T., Chode, S., Parimi, N., Sinacore, D.R., Hilton, T., Armamento-Villareal, R., Napoli, N., Qualls, C., Shah, K., 2011. Weight loss, exercise, or both and physical function in obese older adults. N Engl J Med 364, 1218-1229.

Walston, J., Buta, B., Xue, Q.-L., 2018. Frailty Screening and Interventions: Considerations for Clinical Practice. Clin Geriatr Med 34, 25-38.

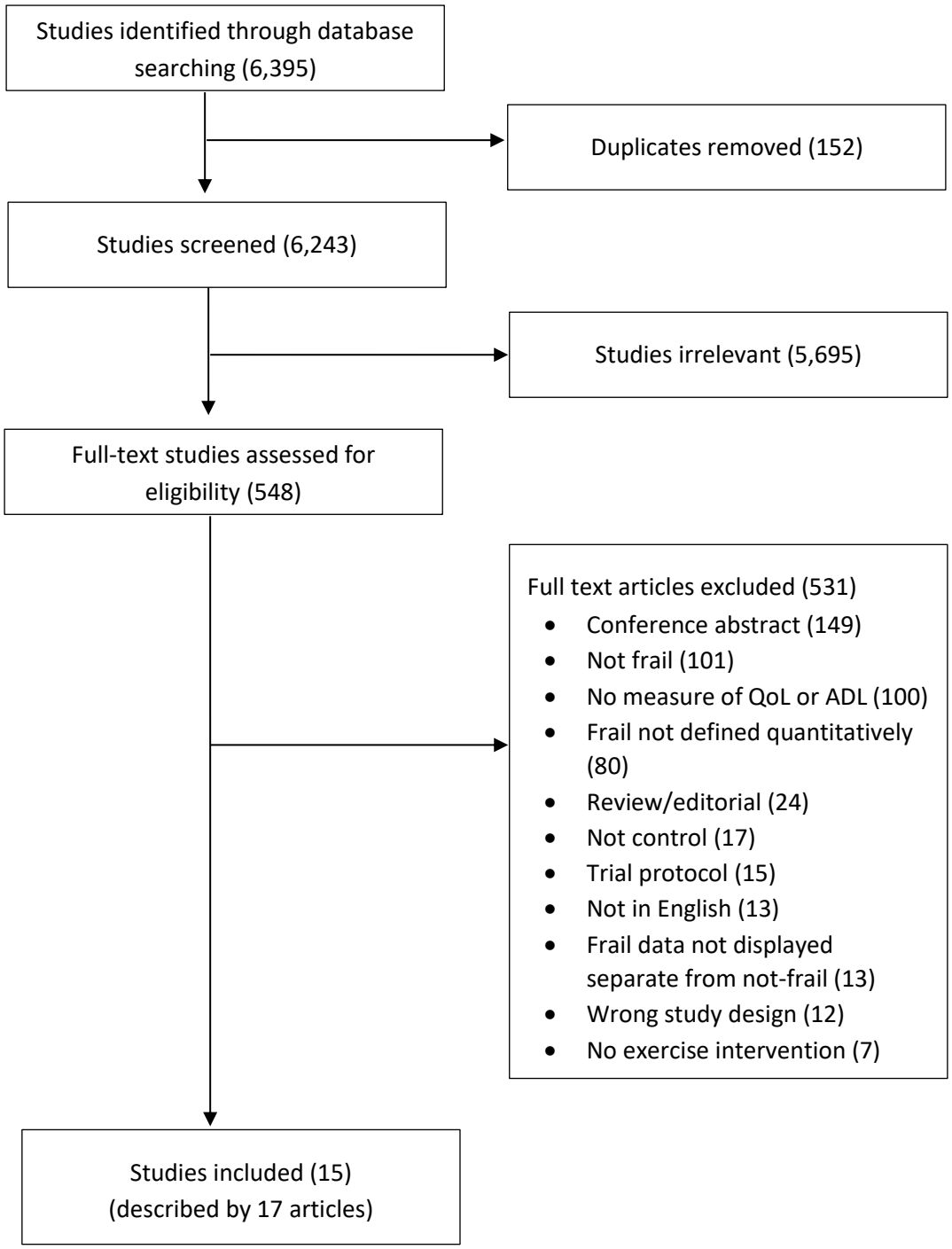


Figure 1. PRISMA flow diagram

470 **Table 1. PEDro Score**

471

	Criteria											
Study	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total score**
EL-Khoury 2015	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Cameron 2013*	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Gine-Gerriga 2013	✓	✓	X	✓	X	X	✓	X	✓	✓	✓	6
Gronstedt 2013	✓	✓	✓	✓	X	X	✓	X	✓	✓	✓	7
Helbostad 2004	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Hoogeboom 2010	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Langlois 2013	X	✓	X	✓	X	X	X	X	X	✓	X	3
Latham 2003	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Littbrand 2009	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Rydwik 2010	✓	X	X	✓	X	X	X	X	✓	✓	X	3
Talley 2017	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Tarazona-Santababina 2016	✓	✓	X	✓	X	X	✓	X	✓	✓	✓	6
Villareal 2011[+]	✓	✓	X	✓	✓	X	X	✓	✓	✓	✓	7
Villareal 2006	✓	✓	✓	✓	X	X	✓	✓	✓	✓	✓	8
Villareal 2017	✓	✓	X	✓	✓	X	X	✓	✓	✓	✓	7

472 *Cameron et al. (Cameron et al., 2013) and Fairhall et al. (Fairhall et al., 2012) describe the same study. Only Cameron et al.
473 is described here because it had a higher PEDro score.
474 +Villareal et al. (Villareal et al., 2011)and Napoli et al. (Napoli et al., 2014)describe the same study. Only Villareal et al. is
475 described here because it had the higher PEDro score.
476 1: *Eligibility criteria were specified (no point awarded for this criterion), 2: Subjects were randomly allocated to groups, 3:*
477 *Allocation was concealed, 4: groups were similar at baseline regarding the most important prognostic indicators, 5: There*
478 *was blinding of all subjects, 6: There was blinding of all therapists who administered the therapy, 7: There was blinding of*
479 *all assessors who measured at least one key outcome, 8: Measures of at least one key outcome were obtained from more*
480 *than 85% of the subjects initially allocated to groups, 9: All subjects for whom outcome measures were available received*
481 *the treatment or control condition as allocated or. where this was not the case. Data for at least one key outcome was*
482 *analysed by "intention to treat", 10: The results of between-group statistical comparison are reported for at least one key*
483 *outcome, 11: The study provides both point measures and measures of variability for at least one key outcome.*
484

485 **Table 2. Summary of studies**

Study	Participants Age (years)		ITT Power Power QoL/ADL Drop out*	Length of trail/ Follow-up	Frailty measure	Type of intervention	Type of exercise	ADL/QOL measure	Overview results: maintenance of decreasing at the main time point	Overview of physical outcome measures at the main time point
	I	C								
EL-Khoury 2015	352 79.8 (2.8)	354 79.6 (2.8)	Yes Yes No 15%	2 years No	Time to walk 6 m >7,<12 sec OR unable to do 4 consecutive tandem stands	Progressive exercise class focusing on fall risk.	Resistance, balance, core, stretching	SF-36	BG: exer ↑physical function sub-scale	Intervention group ↑ physical function.
Cameron 2013, Fairhall 2012	120 83.4 (5.8)	121 83.2 (5.9)	Yes Yes No 20%	12 months No	Fried criteria^	Multi-disciplinary intervention by dietitian, goal setting, physio, geriatrician and rehab physician.	Functional targeting mobility	NEADL, Barthel index, EQ-5D	BG: → Barthel index, EQ-5D	Intervention group ↑ walking speed.
Gine-Gerriga 2013	26 83.9 (2.8)	25 84.1 (3.0)	Yes No No 20%	3 months 6 months	>10 sec to walk 6m OR unable to 5 sit to stands unaided OR frail by the exhaustion criterion from CHS.	Functional circuit training.	Resistance, balance	SF-12	BG exer ↑ SF-12	Intervention group ↑ physical function.
Grönstedt 2013	170 85.0 (7.7)	152 84.9 (7.6)	Yes Yes Yes 17.4%	3 months No	Dependent in ADL by the FIM	Multi-component intervention focusing on the goals of the patient. It could be physical cognitive or social involves goal setting.	Resistance, balance, aerobic, walking	FIM	BG: → FIM WG: exer → FIM, cont ↓ FIM	Intervention group ↑ balance and physical activity levels.
Helbostad 2004	39 81.3 (4.7)	38 80.9 (4.3)	Yes Yes No 14%	3 months 9 months	Fall in the past year OR use of walking aid	Supervised versus home training exercise.	Progressive resistance, progressive balance	SF-36	BG: exer ↑Physical function, exer ↑ role emotional	Both groups ↑ their walking speed and fast walking speed. CT ↑ physical function.
Hoogeboom 2010	10 77.0 (3.0)	11 75.0 (5.0)	Yes No No 5%	3 to 6 weeks No	Clinical Frailty Scale^	Exercise versus no-exercise preoperation.	Resistance, aerobic, functional	HOOD QoL	BG: → HOOD QoL	Intervention group ↑ the chair-rise time and timed-up-and-go.
Langlois 2013	36 71.6 (6.25)	36 73.2 (5.14)	Yes No No 0%	3 months No	≥ 2 of: frail by Fried criteria^, score ≤28/36 on PPT, frail by	Physical exercise-training programme of 12 weeks of 1-hour	Resistance, balance, aerobic	QLSI	BG: exer ↑ QLSI	Intervention group ↑ physical capacity (both functional and physical endurance).

					geriatrician's judgment.	exercise session 3 days a week.				
Latham 2003	120 80 (79-81)	123 78 (77-80)	Yes Yes Yes 9%	10 weeks 3 and 6 months	Winograd Criteria^	Resistance exercise and vitamin D supplementation.	Resistance exercise of a quadriceps programme	SF-36, Barthel Index	BG: → SF-36, Barthel index	Control group had ↑ mobility (timed-up-an-go) than the intervention group.
Littbrand 2009	91 85.3 (6.1)	100 84.2 (6.8)	Yes No No 6%	3 months 6 months	Dependent on one or more ADL by Katz Index	Exercise programme in groups of three to nine participants supervised by two personal trainers (intervention). Occupational therapist supervised the control activities (activities while sitting).	Lower limb resistance, balance, gait, functional	Barthel Index	BG: → Barthel index	Low proportion of participants in the intervention group had ↓ indoor mobility than the control group.
Rydwik 2010	48 83.3 (3.85)	23 82.9 (4.0)	Yes No No 18%	12 weeks 9 and 24 months	Unintentional weight loss >5% in the past year AND/OR BMI <20kg/m^2	Exercise programme or the combination of an exercise programme and nutrition treatment.	Resistance, balance, aerobic, functional	FIM, IAM	BG: → FIM, IAM	Intervention groups ↑ habitual physical activity levels and walking duration.
Talley 2017	23 84.9 (6.4)	19 84.9 (6.4)$^{\$}$	N/A No No 0%	12 weeks No	One of: >3 in Vulnerable Elders Survey, OR gait speed < 0.8 m/s OR Use of walking aid	Combination of bladder intervention and physical activity components.	Resistance, walking	IIQ, UDI	BG: → IIQ, UDI WG: both groups ↑ IIQ, UDI	Intervention groups ↑ more their SPPB balance score and gait score while control group ↑ more their SPPB total score and chair stand score.
Tarazona-Santababina 2016	51 79.7 (3.6)	49 80.3 (3.7)	Yes Yes No 20%	24 weeks No	Fried Criteria^ and Edmonton Frailty Scale^	Multi-component exercise class.	Resistance, balance, aerobic, stretching	Barthel Index, EQ-5D, Lawton Index	BG: exer ↑ Barthel index BG: exer ↑ Lawton Index BG: → EQ-5D	Intervention group ↑ SPPB and physical performance test.
Villareal 2011, Napoli 2014	54 70.0 (4.0)	27 69.0 (4.0)	Yes Yes No 15%	52 weeks No	At least two of: 18 to 31 modified PPT, peak oxygen consumption of 11 to 18 ml/kg; difficulty 2 instrumental ADL or 1	Dietary and exercise intervention.	Resistance, balance, aerobic, stretching	SF-36, IWQOL	BG: exer ↑ FSQ BG: exer ↑ IWQOL	Intervention groups ↑ PPT than control group. PPT ↑ more in the combination group (diet+exercise) than in the diet or exercise groups.

					basic ADL by FSQ. (obese frailty)					
Villareal 2006	17 69.4 (4.6)	10 71.1 (5.1)	Yes Yes No 9%	6 months No	At least two of: 18 to 31 modified PPT, peak oxygen consumption of 11 to 18 ml/kg; difficulty 2 instrumental ADL or 1 basic ADL. (obese frailty)	Diet and exercise therapy.	Resistance, balance, aerobic, stretching	SF-36	BG: exer ↑ SF-36	Intervention group ↑ physical performance and functional status.
Villareal 2017	120 70.0 (4.7)	40 70.0 (5.0)	Yes Yes No 12%	26 weeks No	18 to 31 modified PPT (obese frailty)	Weight management in 3 group exercise (aerobic, resistance and combination) plus one control group.	Resistance, aerobic, and combined	SF-36	BG exer ↑ physical function sub-scale,	Intervention groups ↑ PPT than the control group. PPT ↑ more in the combination group than in the aerobic or resistance groups.

*drop out percentage calculated at main time point.

$ age same between intervention and control groups as age only reported for the whole sample.

^ Indicates that a clinically validated measure of frailty was used.

ADL: activities of daily living; **BG**: differences between groups; **cont**: control group; **BMI:** Body Mass Index; **EQ-5D:** European Quality of Life-5 Dimensions; **FSQ:** Functional Status Questionnaire; **HOOD QoL**: Hip disability and Osteoarthritis Outcome Score Quality of Life; **IAM**: Instrumental Activity Measures; **IIQ**: Incontinence Impact Questionnaire; **ITT:** intention to treat; **IWQOL**: Impact of Weight on Quality of Life; **NEADL:** Nottingham Extended ADL; **PPT**: Physical Performance Test; **QLSI:** Quality of Life Systemic Inventory Questionnaire; **SF-36**: Short-Form 36 Questionnaire; **SPPB:** Short Physical Performance Battery; **UDI:** Urinary Distress Inventory; **WG**: differences within groups.

494 **Appendix**

495 **Searches for frailty review**

496 *a) Embase via OVID*

497 1. frail elderly/

498 2. Frail*.tw.

499 3. or/1-2

500 4. exp exercise/

501 5. muscle strength/ or endurance/ or fitness/ or body equilibrium/

502 6. (exercis* or train* or strength*).tw.

503 7. or/4-6

504 8. 3 and 7

505 9. Case Reports.pt.

506 10. Letter.pt.

507 11. Editorial.pt.

508 12. or/9-11

509 13. 8 not 12

510 14. Human/

511 15. 13 and 1

512 *b) Medline via OVID*

513 1. Frail Elderly/

514 2. frail*.tw.

515 3. or/1-2

516 4. exp Exercise/

517 5. exp muscle strength/ or exp physical endurance/ or physical exertion/ or exp physical fitness/ or
518 postural balance/

519 6. (exercis* or train* or strength*).tw.

520 7. or/4-6

521 8. 3 and 7

522 9. Case Reports/

523 10. Letter/

524 11. Historical article/

525 12. Comment/

526 13. Editorial/

527 14. or/9-13

528 15. 8 not 14

529 16. Humans/

530 17. 15 and 16

531

532 **Table A.1. Effect of exercise in quality of life**

Study	Exercise prescription	ADL/QoL Measure	Effect in quality of life
EL-Khoury 2015	Not specified	SF-36	**SF-36 (mean, SD)** *Physical function* Mean at baseline int: 59.40 ± 21.06; Control; 57.47 ± 20.72. Mean difference at 12th month: -5.27, 95% CI (-9.06 to --1.48), p=0.01. Mean difference at 24th month: -3.90, 95% CI (-7.51 to - 0.30), p=0.03. *Mental health* Mean at baseline Int: 61.33 ± 18.1; Control: 60.70 ± 18.24. Mean difference at 12th month: −2.55, 95% CI (−5.33 to 0.25), p=0.07. Mean difference at 24th month: −1.58, 95% CI (−4.26 to 1.11), p=0.25. *General health* Mean at baseline Int: 57.05 ± 15.80; Control: 54.72 ± 16.05 Mean difference at 12th month: −2.33, 95% CI (−4.86 to 0.20), p=0.04. Mean difference at 24th month: −1.64,95% CI (−4.21 to 0.93), p=0.21. *Vitality* Mean at baseline Int: 48.44 ± 16.26; Control: 46.62 ± 16.15. Mean difference at 12th month: −3.17, 95% CI (−5.71 to −0.62), p=0.01. Mean difference at 24th month: −2.57, 95% CI (−5.21 to 0.08), p=0.06.
Cameron 2013, Fairhall 2012	3-5 x week for one year. Weight-bearing exercise for balance regularly modified	NEADL	**NEADL (mean, SD)** Int. baseline: 9.4 ± 4.13, 3rd month: 10.6 ± 4.6, and 12th month: 10.1 ± 5.2. Control baseline: 9.1 ± 4.3, 3rd month: 9.8 ± 4.7, 12th month: 9.5 ± 4.8. **Barthel index** Int: base: 93.9 ±11.1, 3 month: 108, 94.2 ±11.2, 12 month: 106, 89.5 ±17.5 Cint: base; 92.5 ±14.3, 3 month 117, 93.2 ±13.9, 12 month 108, 86.1 ±24.7 3month p=0.57. 12 month p=0.79 **EQ5D VAS** Int: base; 58.2 ±15.8, 3 month 108, 60.6 ±20.1, 12 month 107, 57.5 ±20.8 Cont: base: 57.89 SD18.4, 3 month 117, 60.3 ±16.9, 12 month 108, 57.7 ±19.7 3 motnh p=0.99, 12 month p =0.91

Gine-Gerriga 2013	2 x week, 12 weeks, alternate days between functional strength and functional balance, the intensity was 1-2 sets of 6-8 reps with a max of 15. once 15 reached then ankle weight progressed y 0.5 kg to a max of 2 kg.	SF-12	**SF-36 (mean, SD)** *Physical function* Int. baseline 33.22 ± 9.49, 12 weeks: 40.8 ± 6.95, 24 weeks: 40.3 ± 7.37 Control baseline: 34.38 ± 10.93, 12 weeks 28.24 ± 6.49, 24 weeks 29.47 ± 5.93. *Physical composite* Int. Baseline: 30.34 ± 6.32, 12 weeks: 35.59 ± 4.41, 24 weeks 36.52 ± 4.47. Control baseline: 32.46 ± 5.42, 12 weeks: 29.80 ± 3.74, 24 weeks: 29.26 ± 3.05. *Mental composite* Int. Baseline: 33.18 ± 7.96, 12 weeks: 38.37 ± 7.14, 24 weeks: 33.94 ± 6.18. Control baseline: 31.11 ± 6.26, 12 weeks: 31.14 ± 8.56, 24 weeks: 30.53 ± 7.41.
Grönstedt 2013	Individual sessions which were progressive not told the frequency or duration of exercise	FIM	**FIM (mean, SD)** *mean change between groups* Intervention: –0.56 ± 7.5 Control: –1.64 ± 10.1 *Mean change within groups* Int. baseline: 48, 95% CI (13–90), 3rd month: 48 95% CI (13 to 91), p=0.832 Control baseline: 47, 95% CI: (13–91), 3rd month: 42 (13 to 91), p=0.012
Helbostad 2004	Class; 2 x week 60 min warm-up 10 min, 20 min strength training, 20 min balance, 10 min cooldown HEP 2x daily	SF-36	**SF-36 (mean. SD)** *Physical function* Int. baseline: 54 ± 23. 3rd month: 61 ± 24. 9th month: 57 ± 24. Control baseline: 47 ± 19. 3rd month: 49 ± 23. 9th month: 47 ± 24. P-value differences at 3rd month= 0.07. 9th month= 0.22 *Role physical* Int. baseline: 49 ± 37. 3rd month: 67 ± 41. 9th month: 51 ± 43. Control baseline: 55 ± 30. 3rd month: 60 ± 36. 9th month: 69 ± 43. P-value differences at 3rd month= 0.17. 9th month= 0.10 *Bodily pain* Int. baseline: 66 ± 29. 3rd month: 70 ± 26. 9th month: 69 ± 29. Control baseline: 64 ± 30. 3rd month: 69 ± 27. 9th month: 67 ± 27. P-value differences at 3rd month= 0.98. 9th month= 0.93 *General Health* Int. baseline: 62 ± 24. 3rd month: 66 ± 25. 9th month: 60 ± 28. Control baseline: 56 ± 23. 3rd month: 58 ± 23. 9th month: 58 ± 24.

			P-value differences at 3rd month= 0.29. 9th month= 0.36 *Vitality* Int. baseline: 44 ± 19. 3rd month: 48 ± 18. 9th month: 44 ± 19. Control baseline: 45 ± 24. 3rd month: 49 ± 20. 9th month: 47 ± 20. P-value differences at 3rd month= 0.47. 9th month= 0.57 *Role emotional* Int. baseline: 66 ± 40. 3rd month: 84 ± 32. 9th month: 83 ± 33. Control baseline: 79 ± 33. 3rd month: 72 ± 39. 9th month: 82 ± 37. P-value differences at 3rd month= 0.003. 9th month= 0.16 *Mental Health* Int. baseline: 74 ± 17. 3rd month: 80 ± 15. 9th month: 15 ± 14. Control baseline: 73± 18. 3rd month: 75 ± 14. 9th month: 72 ± 15. P-value differences at 3rd month= 0.10. 9th month= 0.26
Hoogeboom 2010	min 2 x week 60 min class 5 min warm-up, leg reps 10-20 reps 20-30 min cycle functional ex RPE 13-14 borg scale	HOOD QoL	**HOOD QoL (mean. SD)** Int. baseline: 30.7 ±13.6 Control baseline: 41.0 ± 18.3 Int. after: 36.3 ± 15.8 Control after: 43.3 ± 15.4 Difference between groups: -1.2. 95% CI (-13.0 to 10.7)
Langlois 2013	1 hr, 3x week 12 weeks	The Quality of Life Systemic Inventory questionnaire	The effect in quality of Life Main group effect F(11. 58) = 2.04. p = .04. Int. vs control: global F(1. 68) = 3.97. p = .05 Leisure activities: F(1.68) = 9.13. p = .004. perception of physical capacity. F(1. 68) = 5.76. p = .019. social/family relationships. F(1. 68) = 4.41. p = .039. and physical health. F(1. 68) = 4.40. p = .040.
Latham 2003	3 x week, 10 weeks, intensity at midpoint target 60-80% 1RM	SF-36 Barthel Index	**SF-36 (mean. 95%CI)** *At 3rd month (only in physical function)* Int.: 34. 95%CI (32 to 36); Control: 35. 95% CI (33 to 37). p>0.05. At 6th month Physical component. Int.: 35. 95%CI (33 to 37); Control: 37. 95%CI (35 to 39). p>0.05.

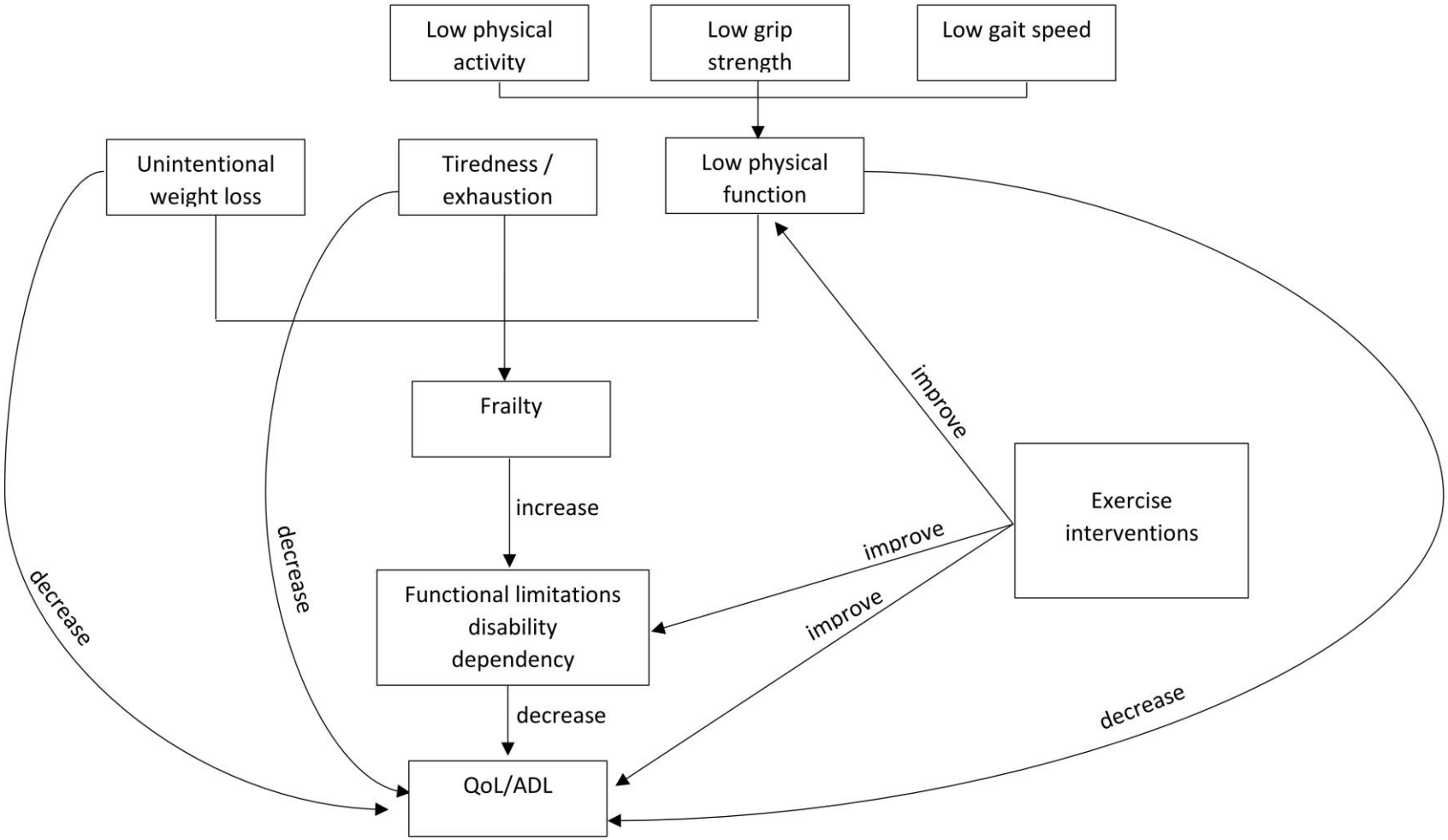
			General health. Int.: 58. 95%CI (53 to 62); Control: 64. 95%CI (59 to 68). p<0.05. Vitality. Int.: 48. 95%CI (44 to 51); Control: 53. 95%CI (49 to 57). p<0.05. **Barthel Index (mean. 95%CI)** Int. baseline: 19. 95%CI (18 to 19). 6th month: 19. 95%CI (19 to 20). Control baseline: 19. 95%CI (18 to 19). 6th month: 20 (19 to 20).
Littbrand 2009	45 min class, five each 2wks over three months (13 weeks), 8-12 RM, progressed as able After the intervention, it was given HEP of functional ex eg squats, walking,	Barthel Index	**Barthel Index (mean. SD)** Int. baseline: 12.8 ±4.5. 3rd month: 12.95 ± 0.27. 6th month: 12.8 ± 0.32 Control baseline: 13.4 ± 3.8. 3rd month: 12.39 ± 0.26. 6th month: 12.10 ± 0.30.
Rydwik 2010	60 min class 2 x week x 12 weeks: 20 min aerobic, resistance training 60-80% max int, balance training Followed by six months of HEP: functional resistance training and balance, + regular walks No intensity given	FIM and IAM	**FIM (median. interquartile range)** Int. (Exercise) baseline: 88 (83–89). 3rd month: 88 (84–89). 9th month: 87 (84–89). 24th month: 84 (77–88). Int. (exercise and nutrition) baseline: 84 (81–88). 3rd month: 87 (83–88). 9th month: 87 (84–89). 24th month: 86 (79–86). Control baseline: 88 (83–90). 3rd month: 88 (85–90). 9th month:90 (87–90). 24th month:88 (84–89). **IAM (median. interquartile range)** Int. (Exercise) baseline: 38 (31–43). 3rd month: 37 (31–45). 9th month: 34 (30–45). 24th month: 34 (23–43). Int. (exercise and nutrition) baseline: 42 (28–48). 3rd month: 44 (35–50). 9th month: 46 (40–48). 24th month: 40 (24–44). Control baseline: 43 (37–48). 3rd month: 44 (37–49). 9th month: 43 (37–53). 24th month: 40 (34–47).
Talley 2017	150 min of moderate-intensity walking a week: Walking 30 min mod int five days a week Strength: 60 min class 2 x week. 10 ex 1 set of each 12-15 reps moderate intensity	IIQ and UDI	**IIQ (mean. SD)** Int. baseline: 45.8 ± 48.8. 3rd month: 39.5 ± 31.6. Control baseline: 58.8 ± 58.8. 3rd month: 40.8 ± 31.6. *p*=0.898 d=-1.3 **UDI (mean. SD)** Int. baseline: 64.8 ± 46.7. 3rd month: 44.0 ± 35.2. Control baseline: 73.7 ± 44.5. 3rd month: 52.2 ± 35.3. *p*=0.897. d=3.3

Tarazona-Santababina 2016	65 min x 5x week, x 24 weeks (proprioception and balance exercises (10e15 minutes), aerobic training (initially at 40% of maximum heart rate increasing progressively to 65%),14 strength training (initially at 25% of 1 repetition maximum to 75%))	Barthel Index. Lawton Index. and EQ-5D	**Barthel Index (mean. SD)** Int. baseline: 88.2 ± 10.9. 6th month: 91.6 ± 0.8. Control baseline: 88.3 ± 10.5. 6th month: 82.0 ± 11.0 p<0.001 (comparison post intervention between groups) Lawton Index (mean. SD) Int. baseline: 6.7 ± 1.1. 6th month: 6.9 ± 0.9. Control baseline: 6.8 ± 1.8. 6th month: 5.37 ± 2.0 p=0.001 (comparison post intervention between groups) **EQ-5D (mean. SD)** Int. baseline: 7.4 ± 2.0. 6th month: 8.2 ± 1.6. Control baseline: 7.7 ± 1.8. 6th month: 7.6 ± 1.3; p = 0.45 (comparison post intervention between groups)
Villareal 2011, Napoli 2014	90 min class 3 per week, aerobic and Resistance: start at 65% of max capability and progressed to 85% of max capability Resistance: start 1-2 sets 8-12 reps, progres to 2-3 sets 6- 8 reps.	SF-36	**SF-36 (mean. SD)** *Physical score* Control baseline: 70.7 ± 12.2. change 6th month: -2.9 ± 13.1. change year: -4.1 ± 10.2. Exercise baseline: 69.3 ± 14.4. change 6th month: 5.7 ± 12.6. change year: 5.7 ± 8.0. Diet+ exercise baseline: 69.8 ± 17.6. change 6th month: 5.1 ± 10.9. change year: 8.6 ± 9.3. *Mental score* Control baseline: 77.2 ± 12.2. change 6th month: 0.4 ± 9.8. change year: -0.5 ± 10.5. Exercise baseline: 75.0 ± 19.7. change 6th month: 2.9 ± 11.6. change year: 5.6 ± 18.3. Diet+ exercise baseline: 79.0 ± 11.0. change 6th month: 4.1 ± 9.2. change year: 3.3 ± 9.9. **Functional Status Questionnaire** Diet–exercise group an increase of 2.7±2.6 points [a 10% change from baseline]. Diet group: 1.3±1.5 points [a 4% change]. **IWQOL** Diet (7.6 6 1.6), exercise (10.1 6 1.6) Diet-exercise (14.0 6 1.4) Control group (0.3 6 1.6) (P = 0.0001–0.03)
Villareal 2006	Three non-consecutives days of exercise: each session lasted 90 minutes and began with 15 minutes of warm-up flexibility exercise followed by 30 min of endurance	SF-36	**SF-36 (mean. SD)** *Physical function* Int. baseline: 60 ± 21.06th month: 83.2 ± 13.9 Control baseline: 67.0 ± 15.1, 6th month: 69.5 ± 22.1. P-value difference: 0.03

	exercise, 30 minutes of strength and 15 minutes of balance.		*Role physical* Int. baseline: 54.4 ± 43.5. 6th month: 78.0 ± 36.3. Control baseline: 62.5 ± 44.5. 9th month: 67.5 ± 42.6. P-value difference: 0.03 *Bodily pain* Int. baseline: 63.3 ± 19.3. 6th month: 73.8 ± 21.4. Control baseline: 74.7 ± 71.0. 6th month: 71.0 ± 18.5. P-value difference: 0.02 *Vitality* Int. baseline: 44.1 ± 17.5. 6th month: 56.3 ± 11.2. Control baseline: 45.4 ± 16.5. 6th month: 47.3 ± 25.5. P-value difference: 0.04 *Social functioning* Int. baseline: 93.0 ± 15.7. 6th month: 92.2 ± 15.7. Control baseline: 93.0 ± 15.6. 6th month: 77.8 ± 22.8. P-value difference: 0.11 *Mental Health* Int. baseline: 85.5 ± 8.1. 6th month: 87.9 ± 10.7. Control baseline: 75.6 ± 20.9. 6th month: 77.2 ± 24.3. P-value difference:0.93 *Role emotional* Int. baseline: 84.1 ± 34.1. 6th month: 98.0 ± 8.1. Control baseline: 81.7 ± 31.6. 6th month: 75.0 ± 40.8. P-value difference: 0.12 *Change in health* Int. baseline: 38.2 ± 12.3. 6th month: 65.7 ± 12.2. Control baseline: 38.0 ± 6.3. 6th month: 38.0 ± 11.4. P-value difference: <0.001
Villareal 2017	**Aerobic group.** Sessions of 60 minutes: 10 minutes of flexibility exercises, 40 minutes of aerobic and 10 minutes of balance. Exercises treadmill walking, stationary	SF-36	**SF-36 (mean. SD)** *Physical score* Control baseline: 47.0 ± 1.7. change 6th month: -1.6 ± 0.8 Aerobic baseline: 48.6 ± 1.4. change 6th month: 6.5 ± 0.7.

	cycling and stair climbing. **Resistance group.** Sessions of 60 minutes: 10 minutes of flexibility exercises, 40 minutes of resistance and 10 minutes of balance. Resistance: nine upper-body and lower-body exercises using weight-lifting machines. Initial sessions were 1 to 2 sets of 8 to 12 repetitions. The combined group participated in the same weight management programme, as well as combined aerobic, and resistance exercise 3 time a week, sessions were 75 to 90 minutes longer. 10 minutes of flexibility exercises, 30 to 40 minutes of aerobic, followed by 30 to 40 minutes of resistance and 10 minutes of balance.		Resistance baseline: 51.0 ± 1.5. change 6th month: 7.4 ± 0.8. Combination baseline: 45.9 ± 1.6. change 6th month: 9.5 ± 0.7. *Mental score* Control baseline: 42.7 ± 0.7. change 6th month: -0.7 ± 0.6. Aerobic baseline: 43.4 ± 0.9. change 6th month: 1.9 ± 0.5 Resistance baseline: 42.7 ± 0.9. change 6th month: 1.9 ± 0.6 Combination baseline: 45.1 ± 0.9. change 6th month: 2.6 ± 0.5.

535 **Figure A.1. Pathway of interaction of exercise interventions, physical function and QoL and ADLs**



536