

# The Effect of Therapeutic Exercise Interventions on Physical and Psychosocial Outcomes in Adults Aged 80 Years and Older: A Systematic Review and Meta-Analysis

Philippa J.A. Nicolson, Vicky Duong, Esther Williamson, Sally Hopewell, and Sarah E. Lamb

This systematic review aimed to evaluate the effects of therapeutic exercise on physical and psychosocial outcomes in community-dwelling adults aged 80 years or older. Databases were searched from inception to July 8, 2020. Randomized controlled trials (RCTs) were screened by two reviewers who extracted data and assessed study quality. Sixteen RCTs (1,660 participants) were included. Compared to nonexercise controls there was no evidence of an effect of exercise on performance based (standardized mean differences: 0.58, 95% confidence interval: [-0.19, 1.36];  $I^2$ : 89%; six RCTs; 290 participants; very low-quality evidence) or self-reported physical function (standardized mean differences: 1.35, 95% confidence interval: [-0.78, 3.48];  $I^2$ : 96%; three RCTs; 280 participants; very low-quality evidence) at short-medium term follow-up. Four RCTs reporting psychosocial outcomes could not be combined in meta-analysis and reported varying results. Exercise appeared to reduce the risk of mortality during follow-up (risk ratio: 0.47, 95% confidence interval: [0.32, 0.70];  $I^2$ : 0.0%; six RCTs; 1,222 participants; low-quality evidence).

**Keywords:** age 80+, exercise therapy, function, quality of life

The global population is progressively aging. Approximately 3 million people in the United Kingdom were aged 80 years or older in 2018, and this group is projected to increase to almost 6 million by 2043, making it the fastest growing population group (Office for National Statistics, 2018). Adults aged 80 years or older are the least physically active, and have the highest health care expenditure (England, 2018; Manini & Pahor, 2009). More than 85% of people aged 80 years or older in the United Kingdom reside in the community, rather than in nursing homes (Office for National Statistics, 2012). Optimizing physical function, quality of life, and psychosocial outcomes among this group is essential to facilitate ongoing independence.

Therapeutic exercise is participation in physical activity that is planned, structured, repetitive, and purposeful for the improvement or maintenance of a specific health condition (World Health Organization, 2010). The benefits of therapeutic exercise for all adults, and for many age-related conditions such as osteoarthritis and frailty, are well established (Fransen et al., 2015; Fransen, McConnell, Hernandez-Molina, & Reichenbach, 2014; Heyn,

Johnson, & Kramer, 2008; Silva, Aldoradin-Cabeza, Eslick, Phu, & Duque, 2017). However, adults aged 80 years and older are significantly underrepresented in existing trials, and findings among adults aged in their 60s and 70s cannot necessarily be applied to those aged 80 years or older (Izquierdo, Morley, & Lucia, 2020; Witham et al., 2020).

Adults aged 80 years or older are a more heterogeneous population than younger older adults, with significant variability in magnitude of multimorbidity, disability, and physical and cognitive function (Collerton et al., 2016; Lafortune, Béland, Bergman, & Ankri, 2009; Santoni et al., 2015). In addition, loss of muscle strength accelerates from 10% to 15% per decade up to the age of 70 years, to 25% to 40% per decade beyond 70 years of age (Goodpaster et al., 2006; Hughes et al., 2001). Adults aged 80 years and older are often excluded from randomized controlled trials (RCTs) of exercise interventions for older adults, and clinically are less likely to be prescribed exercise than younger people (Smith, Collier, Smith, & Mansfield, 2019; Witham et al., 2020). A recent U.K. research priority setting exercise found that health professionals questioned whether exercise would have any impact on the health of people aged 80 years and older, and people aged 80 years or older raised concerns about the safety of exercise for them (James Lind Alliance, 2018).

To our knowledge no systematic reviews have been published evaluating the effects of therapeutic exercise specifically among community-dwelling adults aged 80 years or older. To inform patients, researchers, clinicians, and stakeholders about the effectiveness of therapeutic exercise for this rapidly growing population group a comprehensive review is required. Findings will also provide knowledge that will support the development of a therapeutic exercise intervention for people aged 80 years or older.

The objective of this systematic review was to evaluate the effects of therapeutic exercise interventions on physical function, health-related quality of life, and psychosocial outcomes in community-dwelling adults aged 80 years or older.

© 2022 The Authors. Published by Human Kinetics, Inc. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0, which permits unrestricted noncommercial and commercial use, distribution, and reproduction in any medium, provided the original work is properly cited, the new use includes a link to the license, and any changes are indicated. See <http://creativecommons.org/licenses/by/4.0>. This license does not cover any third-party material that may appear with permission in the article.

Nicolson, Williamson, Hopewell, and Lamb are with the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom. Duong is with the Department of Rheumatology, Institute of Bone and Joint Research, Kolling Institute, Royal North Shore Hospital and Northern Clinical School, The University of Sydney, St Leonards, NSW, Australia. Lamb is also with the College of Medicine and Health, University of Exeter, Exeter, United Kingdom. Nicolson ([philippa.nicolson@ndorms.ox.ac.uk](mailto:philippa.nicolson@ndorms.ox.ac.uk)) is corresponding author.

## Methods

This review is reported according to the Preferred Reporting Items for Systematic Review and Meta-Analyses guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). The protocol was prospectively registered with the International Prospective Register of Systematic Reviews database (PROSPERO: CRD42020196697).

### Search Strategy

The search strategy was developed in consultation with an academic librarian from the University of Oxford. Three components of the search strategy were developed separately (population, exercise, and RCTs), then combined using database-specific truncation terms. Search terms included controlled vocabulary (e.g., MeSH) and free-text terms. We searched the MEDLINE (via OVID), EMBASE (via OVID), and CINAHL (via EBSCOhost) databases from inception to July 8, 2020. No date or language limits were applied. The search strategy for MEDLINE is presented in [Supplementary Table S1](#) (available online), and this was translated to the relevant syntax for each database. Supplementary searches of reference lists of included studies, relevant systematic reviews and the World Health Organization's International Clinical Trials Registry (<http://apps.who.int/trialsearch>) were also undertaken.

### Study Selection

Randomized controlled trials involving therapeutic exercise for community-dwelling adults aged 80 years or older were eligible. Trials that included younger participants were eligible if the mean age minus 1 *SD* was 80 years or higher, or if there was clear presentation of results by strata for those aged 80 years or older. Trials with mixed community-dwelling and nursing home populations were eligible if <20% of the population resided in nursing homes, or if results were presented separately for community-dwelling participants. Interventions delivered to participants in inpatient hospital settings were not eligible. We excluded studies that only recruited participants affected by the following conditions: Parkinson's disease, Huntingdon's disease, Alzheimer's disease, advanced dementia, or stroke. Due to the nature of these neurological conditions, findings in these populations may not be applicable to adults aged 80 years or older without the specific disease. Some individuals with these (and other) health conditions may be included in studies of the general community; these trials were included.

Therapeutic exercise interventions were classified according to the ProFaNE taxonomy: aerobic (aimed at cardiovascular conditioning), resistance (contracting the muscles against external force such as weights, resistance band, or body weight), functional training (utilizing functional activities such as sit to stand as the training stimulus), balance training (challenging specific aspects of the balance systems), gait training (specific correction of walking technique), flexibility (stretching exercises which are practiced and progressed), or 3D (constant movement in a controlled, fluid, repetitive way through all three spatial dimensions, e.g., Tai Chi; Lamb et al., 2011). Interventions that included more than one type of exercise were classified as multicomponent. We included therapeutic exercise interventions with or without additional interventions (e.g., diet or pharmacotherapy) as long as the cointerventions were delivered to both intervention and control groups. We excluded interventions without a structured exercise program (e.g., providing pedometers without an exercise plan). Exercise could be delivered as a group intervention or individually, in an outpatient clinical setting,

the participant's home or a community location (e.g., community center or gym). Interventions could be delivered in-person or via video consultation. Any comparator treatment, including usual care; no treatment; an alternative exercise treatment; pharmacotherapy, education or nutritional interventions were eligible.

Our primary outcomes were measures of physical function (performance-based or self-reported questionnaire) and health-related quality of life indices. Secondary outcomes were measures of psychosocial health (including anxiety, depression, and loneliness), falls, adverse events, and mortality. Trials were eligible if they reported one or more of these outcomes.

Search results were imported into Covidence (Veritas Health Innovation, Melbourne, VIC, Australia, [www.covidence.org](http://www.covidence.org)) for screening. A two-step process was used for screening and selection. In the first step, titles, and abstracts of all identified RCTs were independently screened by two reviewers (PN and VD). Following title and abstract screening, the full text of all potentially eligible articles was retrieved, and each screened independently for final inclusion by the same two reviewers (PN and VD). Disagreements were resolved through discussion, with an adjudicator (SH) available to address any unresolved disagreements.

### Data Extraction

Data were extracted independently by two reviewers (PN and VD) using a customized piloted data extraction form in Microsoft Excel (Microsoft Corp., WA, DC, 2019). We extracted the following data:

- Trial design (setting, sample size, inclusion and exclusion criteria, method of recruitment, length of follow-up, number of participants randomized, and number analyzed in the intervention, and comparator groups).
- Characteristics of participants (age, gender, ethnicity, number randomized/analyzed, and dropouts in each arm).
- Type of intervention (experimental/control components). Exercise details: type, supervised/unsupervised, group/individual, duration (weeks), frequency (per week), and intensity (subjective: e.g., self-rated scale or objective: e.g., heart rate monitor).
- Adherence to intervention (method of assessment and reported data).
- Outcomes measured (all time points). We extracted information on differences in outcomes between groups at follow-up only. We categorized follow-up as short-term ( $\geq 3$  months postrandomization), intermediate term (4–11 months postrandomization), and long term ( $\geq 12$  months postrandomization). To avoid multiplicity, we selected one outcome measure when multiple outcome measures were reported for the same outcome in a trial. For performance-based physical function we prioritized: (a) the Timed Up and Go Test, (b) Sit to stand test (5× sit to stand or 30-s sit to stand), and (c) Tinetti test. For self-reported physical function we prioritized (a) the Barthel Index and (b) any other self-reported measures of physical function. For health-related quality of life outcomes we prioritized (a) the EuroQol Group 5-Dimension Self-Report Questionnaire, (b) the 36-Item Short Form Survey or the 12-Item Short Form Survey, and (c) any other self-reported health-related quality of life measure. For anxiety symptoms we prioritized (a) the anxiety subscale of the Hospital Anxiety and Depression Scale (anxiety) and (b) any other self-reported anxiety measure. For depression symptoms we prioritized (a) the depression subscale of the Hospital Anxiety and Depression Scale (depression) and (b) any other self-reported depression measure.

- Adverse events, categorized according to the U.S. Food and Drug Administration definition of serious and nonserious (U.S. Food and Drug Administration, 2020), and mortality.

If data were not reported in full in the published manuscript we emailed the corresponding author on two occasions, 1 month apart (if required) requesting the missing data.

## Risk of Bias Assessment and Overall Evaluation of the Quality of the Evidence

Two reviewers (PN and VD) independently assessed the risk of bias of each included trial using the Revised Cochrane Risk of Bias Tool 2.0 (RoB 2; Sterne et al., 2019). Any disagreements in assessment between the two reviewers were discussed until consensus was reached.

We assessed the risk of bias for five domains: bias arising from the randomization process or lack of allocation concealment, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome (blinding), and bias in selection of the reported result. Within each domain, the two reviewers answered one or more signaling questions leading to judgments for each domain as “high risk of bias,” “some concerns,” or “low risk of bias.” The judgments within each domain led to an overall risk of bias judgment (Higgins et al., 2020).

The overall quality of evidence was evaluated using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach (Cochrane Handbook for Systematic Reviews of Interventions, 2019). GRADE is a systematic approach to rate the certainty of evidence (high, moderate, low, or very low) across studies for each outcome of interest. Five domains are assessed, including methodological flaws of the included studies, heterogeneity of results across trials, generalizability of the findings to the target population, precision of the estimates, and risk of publication bias.

## Data Synthesis

Descriptive characteristics of all included trials were summarized in tables and synthesized in narrative format by outcome. We performed meta-analyses using a random effects model as heterogeneity was expected in participant, intervention, and outcome characteristics. For continuous outcomes measured using the same scale we calculated the mean difference (MD) and for continuous outcomes measured using different scales we calculated the standardized mean differences (SMDs) with 95% confidence intervals (CIs). For binary outcomes we calculated the risk ratio (RR). Statistical heterogeneity across pooled studies was quantified using the  $I^2$  statistic (Higgins et al., 2011). Separate meta-analyses were undertaken for studies using nonexercised comparators and trials using a different type of exercise as the comparator. Forest plots were presented by exercise type. Trials with short-medium term follow-up were combined due to the small number of studies, and those with long-term follow-up were analyzed separately. When trials reported multiple time points in one category the longest time point was included in meta-analysis. Meta-analyses were undertaken using Review Manager (Review Manager (RevMan) [Computer program]. Version 5.4, The Cochrane Collaboration, 2020).

## Sensitivity Analysis

Where sufficient trials were identified, sensitivity analyses were planned to assess the effects of the exercise interventions excluding

trials that included participants younger than 80 years of age; that had one or more “high risk” domains of risk of bias. Meta-regression to explore the impact of trial level characteristics on outcomes was also planned where ≥10 studies provided data for each outcome. Due to the limited number of studies included in each of the meta-analyses, meta-regression was not undertaken.

## Results

### Study Selection

The literature search identified 5,232 unique citations. Of these, 46 articles progressed to full-text eligibility review. We included 16 RCTs (reported in 20 articles) in narrative analysis, and 10 of these in quantitative analysis (Figure 1).

### Characteristics of Included Studies

Characteristics and details of the 16 included RCTs (1,660 participants) are presented in Table 1. Studies were published between 1997 and 2020. Trials were conducted across 13 countries, most commonly in Northern Europe (four studies; Bårdstu et al., 2020; Bechshoft et al., 2017; Hvid et al., 2016; Luukinen et al., 2006; Luukinen et al., 2007) or Australia/New Zealand (three studies; Campbell et al., 1997; Hamdorf & Penhall, 1999; Rosie & Taylor, 2007). Three RCTs were conducted in communities specifically developed for older adults (retirement villages or senior communities where older people lived independently; Bonnefoy et al., 2003; de Bruin & Murer, 2007; Hartshorn, Delage, Field, & Olds, 2002). All other RCTs were conducted with people aged 80 years or older who resided independently in the wider community.

Sample sizes of included RCTs varied widely (range: 26–486). The median sample size was 61. Participants had a median age of 84.2 years (interquartile range: 83.4–86.1 years). The 12 RCTs that recruited both male and female participants included a median of 72.4% females (interquartile range: 61.5–82.9%). Three RCTs recruited only females (Campbell et al., 1997; Hamdorf & Penhall, 1999; Kim et al., 2019), and one RCT included only male participants (Kalapotharakos, Diamantopoulos, & Tokmakidis, 2010).

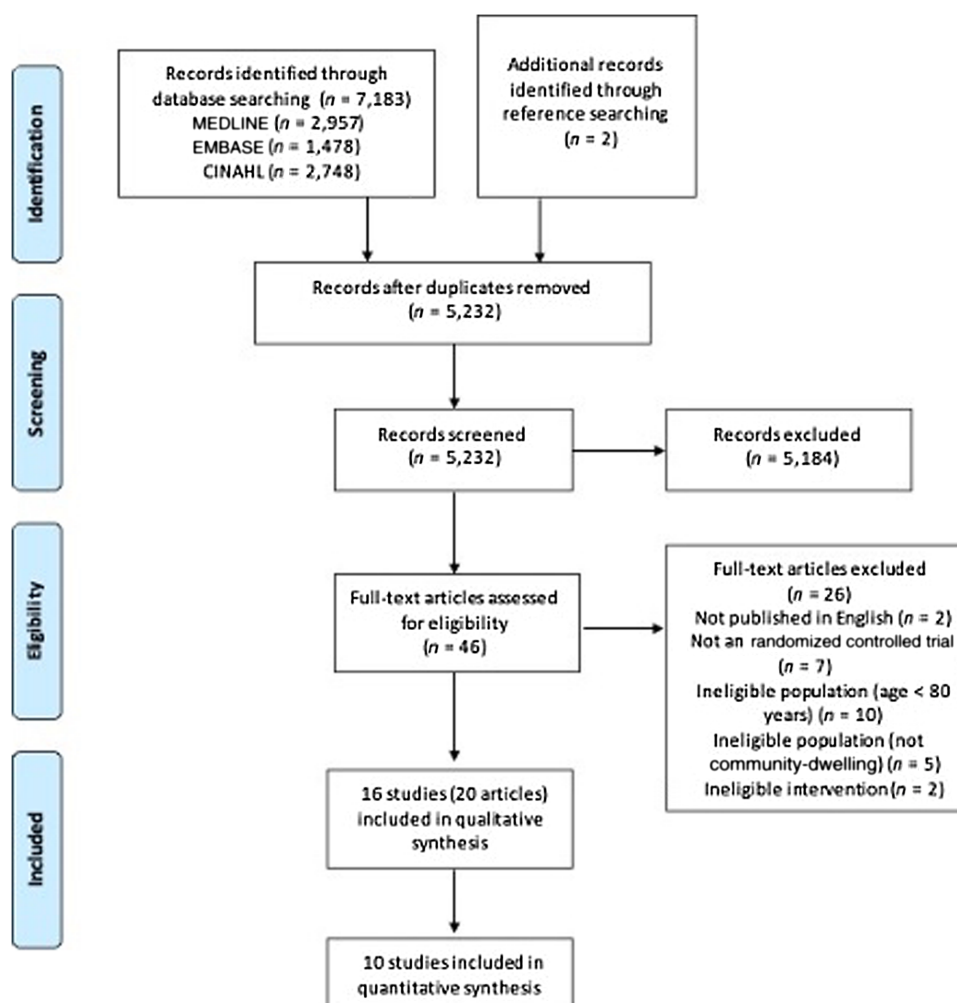
### Risk of Bias

Only one of the included RCTs (Campbell et al., 1997) was deemed low risk of bias overall (Supplementary Table S2 [available online]). Ten of the 16 RCTs (Ansai, Aurichio, Gonçalves, & Rebelatto, 2016; Ansai & Rebelatto, 2015; Bårdstu et al., 2020; Bechshoft et al., 2017; Bonnefoy et al., 2003; Gine-Garriga et al., 2010; Hartshorn et al., 2002; Kim et al., 2019; Luukinen et al., 2006; Rosie & Taylor, 2007; Siemonsma et al., 2018) were deemed as “some concerns” and five (Cancela Carral, Pallin, Orbegozo, & Ayan Perez, 2017; de Bruin & Murer, 2007; Hamdorf & Penhall, 1999; Hvid et al., 2016; Kalapotharakos et al., 2010) were rated as “high” risk. Sources of bias were most commonly Domain 3 (missing outcome data) and Domain 4 (measurement of the outcome).

### Intervention and Comparator Groups

The most common type of exercise intervention was multicomponent ( $n = 7$ , all included resistance training; Ansai et al., 2016; Bonnefoy et al., 2003; Campbell et al., 1997; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Kim et al., 2019; Luukinen et al., 2006), followed by resistance training ( $n = 5$ ; Bårdstu et al., 2020; Bechshoft et al., 2017; Cancela Carral et al., 2017; Hvid et al., 2016; Kalapotharakos





**Figure 1** — Flow chart of study selection process.

et al., 2010), aerobic exercise ( $n=2$ ; Cancela Carral et al., 2017; Hamdorf & Penhall, 1999), functional training ( $n=2$ ; Rosie & Taylor, 2007; Siemonsma et al., 2018), and 3D exercise ( $n=1$ ; Hartshorn et al., 2002). Intervention duration ranged from 2 weeks (Hartshorn et al., 2002) to 24 months (Campbell, Robertson, Gardner, Norton, & Buchner, 1999; median duration: 12 weeks). Intensity of exercise training was reported in nine of the included RCTs (Ansai et al., 2016; Bårdstu et al., 2020; Bechshoft et al., 2017; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Hamdorf & Penhall, 1999; Hvid et al., 2016; Kalapotharakos et al., 2010). Exercise intensities varied, resistance training was most commonly at 70% of one to three repetition maximum, and aerobic exercise was most commonly at 60% of maximum heart rate.

All trials provided partial or total supervision of exercise interventions. Three RCTs included unsupervised home exercises in addition to supervised sessions (Campbell et al., 1997; Luukinen et al., 2006; Rosie & Taylor, 2007). Exercise interventions were delivered in group settings ( $n=8$ ; Bonnefoy et al., 2003; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Hamdorf & Penhall, 1999; Hartshorn et al., 2002; Kalapotharakos et al., 2010; Kim et al., 2019), individually ( $n=7$ ; Ansai et al., 2016; Bårdstu et al., 2020; Bechshoft et al., 2017; Campbell et al., 1997; Hvid et al., 2016; Rosie & Taylor, 2007; Siemonsma et al., 2018), and as a combination of group and

individual sessions ( $n=1$ ; Luukinen et al., 2006). Exercise interventions were performed at community health centers or community facilities ( $n=9$ ; Bårdstu et al., 2020; Bonnefoy et al., 2003; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Hamdorf & Penhall, 1999; Hartshorn et al., 2002; Kalapotharakos et al., 2010; Luukinen et al., 2006), in participants' homes ( $n=4$ ; Bechshoft et al., 2017; Campbell et al., 1997; Rosie & Taylor, 2007; Siemonsma et al., 2018), or at university testing facilities ( $n=2$ ; Ansai et al., 2016; Kim et al., 2019).

Comparator groups varied widely. Nonexercise comparators (13 RCTs) included health education (Gine-Garriga et al., 2010), physical activity counseling (Bårdstu et al., 2020), memory sessions (Bonnefoy et al., 2003), social visits (Campbell et al., 1997), protein supplementation (Bechshoft et al., 2017; Kim et al., 2019), preventative physical therapy (Siemonsma et al., 2018), usual care (Luukinen et al., 2006), and wait list control (Ansai et al., 2016; Hamdorf & Penhall, 1999; Hartshorn et al., 2002; Hvid et al., 2016; Kalapotharakos et al., 2010). Exercise comparator groups were used in five RCTs (Ansai et al., 2016; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Kalapotharakos et al., 2010; Rosie & Taylor, 2007). Two trials compared multicomponent interventions to resistance exercises (Ansai et al., 2016; de Bruin & Murer, 2007), one study used resistance training followed by a detraining period (Kalapotharakos et al., 2010), one trial compared aerobic,

**Table 1** Characteristics of Included Studies

Study	Country	Sample size	Inclusion criteria	Exclusion criteria	Mean age (SD) %Women	Exercise intervention type	Exercise intervention delivery	Exercise intervention frequency/duration	Adherence	Control intervention	Follow-up time point/s	Outcome/s
Ansaï and Rebelatto (2015); Ansaï et al. (2016)	Brazil	69	Aged ≥80 years, community-dwelling, and sedentary, and able to walk alone with/without walking aid	Any absolute CI to exercise; MMSE score <educational level – 1SD	82.4 (2.4) 68.1% female	Multicomponent: aerobic, resistance, and balance exercises	Supervised at university gym Individual	60 min 3×/week 16 weeks	Session attendance Adherent ≥24/48 sessions: 34.7% MT group; and 56.5% RT group	1. Resistance: machine-based exercises 2. No intervention	16 weeks 22 weeks	TUGT motor TUGT cognitive 5× STS GDS Falls
Bårdstu et al. (2020)	Norway	110	Aged ≥70 years, living at home, and receiving home care due to disabilities	Serious cognitive impairment, physical conditions that could affect testing/trainings, and any other CI for training	Exercise 86.5 (range 80–90) 66% female Control 86.0 (80–90) 51% female	Resistance	Supervised at local health centers Individual	30–45 min 2×/week 8 months	Session attendance Mean: 51%	Physical activity counseling	4 months 8 months	5× STS 8 foot up and go Usual gait speed Maximum gait speed Stair climb
Bechshøft et al. (2017)	Denmark	26	Aged ≥83 years; living in one suburb of Copenhagen	Any CI for training or MRI, unstable cardiac arrhythmia, GOLD 3–4, GFR <30 ml/min, diabetes, severe cognitive impairment, systemic corticosteroid, or anticoagulants	86.9 (3.2) 38% female	Resistance + protein supplement	Supervised at home OR in training facilities at local nursing homes Individual	3 × 12 repetitions increasing to 5 × 6 repetitions 3×/week 12 weeks	Session attendance Mean: 32 ± 1/36 sessions	Protein supplement	12 weeks	30-s STS Usual gait speed DMMI
Bonnefoy et al. (2003)	France	57	“Frail older adults;” resident in retirement homes (independent living)	Uncontrolled diseases; dementia, Type 1 diabetes, severe renal insufficiency, long-term corticosteroid therapy, and age <72 years	Exercise 83.0 (1.1) Control 84.0 (1.1) 87.7% female	Multicomponent: resistance and balance and flexibility exercises	Supervised at retirement home Group	60 min 3×/week 9 months	Session attendance Mean: 3 months: 70% 9 months: 63%	Memory sessions	3 months 9 months	5× STS Usual gait speed 6× stair climb
Campbell et al. (1997, 1999)	New Zealand	233	Women, aged ≥80 years, able to move around own home, and not receiving physiotherapy	Deemed unsafe to take part by GP	84.1 (3.3) 100% female 2 years follow-up: 83.9 (3.0) 100% female	Multicomponent: resistance, balance, and functional + walking program	4× home visits + unsupervised home program Individual	30 min 3× week 12 months	Exercise diary calendar (monthly) 12 months: 42% completing exercises ≥3×/week	Social visits + telephone calls	6 months 12 months 24 months	5× STS Physical self-maintenance scale Falls

(continued)

Table 1 (continued)

Study	Country	Sample size	Inclusion criteria	Exclusion criteria	Mean age (SD) %Women	Exercise intervention type	Exercise intervention delivery	Exercise intervention frequency/duration	Adherence	Control intervention	Follow-up time point/s	Outcome/s
Cancela Carral et al. (2017)	Spain	36	Aged ≥80 years, attending day care centers for older adult, and able to follow simple instructions	Unstable cardiovascular, metabolic, neuromuscular conditions, and any chronic disease that could limit training	87.91 (4.70) 80.6% female	1. Aerobic (seated exercises) 2. Resistance (seated exercises)	Supervised at day care centers Group	45 min 3× week 12 weeks	Session attendance 37/50 attended >80% of the sessions	Joint mobility exercises	3 months	Tinetti test Barthel Index TUGT
de Bruin and Murer (2007)	Switzerland	32	Aged ≥70 years, resident in senior's hostel (independent living), and able to walk 6 m	Severe cognitive impairment, rapidly progressive illness, MI, lower extremity fracture within 6/12, insulin-dependent diabetes, and undergoing resistance training	Exercise 85.4 (5.4) 91.7% female Control 86.5 (4.9) 61.5% female	Multicomponent: resistance + functional exercises	Supervised at senior's hostel Group	Resistance: 45 min 2×/week + Function: 30 min 1× week 12 weeks	Session attendance "One participant excluded as they attended only 54% of the sessions" No further detail given	Resistance exercises	12 weeks	Tinetti test 5× STS
Giné-Garriga et al. (2010, 2013)	Spain	51	Aged 80–90 years, classified as frail (physical tests + CES-D or Fried criteria)	Unable to walk, undergoing exercise, severe dementia, stroke, hip fracture, MI, THR, or TKR within 6/12	84 (2.9)	Multicomponent: resistance + functional exercises	Supervised at primary care facilities Group	45 min 2× week 12 weeks	Session attendance 90% completed all 24 exercise sessions	Health education meetings	12 weeks 36 weeks	Barthel Index 5× STS Usual gait speed Fast gait speed TUGT motor cognitive SF-12
Hamdorf and Penhall (1999)	Australia	49	Women, community-dwelling, and not functionally impaired	Participating in formal programs or were regularly active; CI to exercise testing (medical screening)	Exercise 82.4 (SEM: 0.66) Control 83.1 (SEM: 0.69) 100% female	Aerobic (walking)	Supervised in community outdoor spaces Group	5 min in Week 1 to 25 min in Week 22 2× week 6 months	Session attendance Mean: 89.7%	No intervention	6 months	Philadelphia Geriatric Morale scale
Hartshorn et al. (2002)	United States	32	Resident in retirement community	NR	86.0 (3.3) 81.3% female	3D (movement therapy)	Supervised in retirement community Group	50 min 2×/week 2 weeks	NR	Wait list control	2 weeks	Tinetti test POMS

(continued)

Table 1 (continued)

Study	Country	Sample size	Inclusion criteria	Exclusion criteria	Mean age (SD) %Women	Exercise intervention type	Exercise intervention delivery	Exercise frequency/ duration	Adherence	Control intervention	Follow-up time point/s	Outcome/s
Hvid et al. (2016)	Denmark	65	Aged $\geq 75$ years, 3-m usual gait speed $< 0.90$ m/s, and MMSE score $> 21$	Amputation, major physical impairments, terminal disease, surgery or fractures within 6/12, uncontrolled hypertension (BP $> 160/100$ mmHg), and severe pain	Exercise 82.3 (1.3) 56% female Control 81.6 (1.1) 67% female	Resistance (high intensity)	Supervised (location NR) Individual	12 weeks	Session attendance/ completion 80% of participants attended $\geq 20/24$ sessions and completed $\geq 80\%$ of each session	No intervention	12 weeks	Maximum gait speed
Kalapotharakos et al. (2010)	Greece	47	Men, aged $\geq 80$ years, and independent BDL/IADL	Cognitive impairment; unstable chronic disease that would limit exercise	83.4 (2.8) 0% female	Resistance	Supervised at older adults public care centre Group	60 min 2x/week 14 weeks	Session attendance Mean: 8 weeks: 90% 14 weeks: 85%	1. Resistance + detraining period 2. No intervention	12 weeks	TUGT 5x STS 6-min walk
Kim et al. (2019)	Japan	126	Aged $\geq 75$ years, community-dwelling, walking speed $< 1.0$ m/s, step width $> 10$ cm, and stride length $< 100$ cm	Severely impaired mobility or unstable cardiac conditions $< 1.0$ m/s, step width $> 10$ cm, and stride length $< 100$ cm	Ex + protein: 82.8 (2.8) Ex + placebo: 83.1 (3.3) Protein: 82.9 (2.9) Placebo: 83.8 (3.3) 100% female	Multicomponent: resistance, balance + gait retraining + protein supplement	Supervised at university department Group	60 min 2x week 12 weeks	NR	1. Protein supplement 2. Placebo supplement	12 weeks	TUGT Usual walking speed Grip strength
Luukinen et al. (2006, 2007)	Finland	486	$\geq 2$ falls in past 12/12, loneliness, poor self-rated health, depression, low cognitive status, impaired vision/hearing/ balance, slow walking speed, and difficulty standing from chair	NR	88 (3) 21% female	Multicomponent (pragmatic, personalized exercise plan)	Varied Supervised at community centers and/or home based. Group and/or individual	Variable. Home exercises daily. Median duration: 16 months	NR	Routine care	28 months	ADL score Mobility score Self-reported (verified) falls

(continued)

Table 1 (continued)

Study	Country	Sample size	Inclusion criteria	Exclusion criteria	Mean age (SD) %Women	Exercise intervention type	Exercise intervention delivery	Exercise intervention frequency/duration	Adherence	Control intervention	Follow-up time point/s	Outcome/s
Rosie and Taylor (2007)	New Zealand	68	Aged ≥80 years, able to walk 4 m, sedentary, limited "a lot" in >1 activity in SF-36 PF-10	Receiving physiotherapy, medically unstable, and any CI for exercise	85.2 (3.6) 71.2% female	Functional training (repeated STS with biofeedback)	Supervised + unsupervised at home. Individual	Daily 6 weeks	Exercise diary (number of days exercises completed) Mean: 31/42 days 74%	Resistance exercises (seated)	6 weeks	30-s STS 15-s step test Usual gait speed LLFDI Falls Efficacy Scale
Siemonsma et al. (2018)	The Netherlands	155	Aged ≥75 years, MMSE ≥18, speak Dutch	Admitted to nursing home, life expectancy <3 months, serious psychiatric illness, inability to follow instructions, and current physiotherapy	Exercise 84.0 (Q1 79.4; Q3 88.7) 72% female Control 83.9 (80.2; 86.4) 75% female	Functional training	Supervised at home. Individual	Up to 18 sessions 12 weeks	NR	Preventative physical therapy	12 weeks	TUGT 5× STS GARS Modified Katz-15

*Note.* ADL = activities of daily living; BDL/IADL = basic activities of daily living / instrumental activities of daily living; CES-D = Center for Epidemiological Studies-Depression; CI = contraindications; Con = control; DMMI = DeMorton Mobility Index; Ex = exercise; GARS = Groningen Activities Restriction Scale; GDS = Geriatric Depression Scale; GFR = glomerular filtration rate; GOLD = Global Initiative for Chronic Obstructive Lung Disease scale; GP = general practitioner; LLFDI = Lower Limb Functional Disability Index; MI = myocardial infarction; MMSE = Mini-Mental State Examination; MRI = magnetic resonance imaging; MT = multicomponent training; NR = not reported; POMS = profile of mood states; RT = resistance training; SF36 PF-10: short form-36 Questionnaire physical function component; STS = sit to stand; THR = total hip replacement; TKR = total knee replacement; TUGT = Timed Up and Go test; BP = blood pressure; SF-12 = 12-Item Short Form Survey; SF-36 = 36-Item Short Form Survey.



resistance and joint mobility exercise interventions (Cancela Carral et al., 2017), and one compared a functional training program to resistance exercises (Rosie & Taylor, 2007). Two RCTs (Ansai et al., 2016; Kalapotharakos et al., 2010) included both nonexercise and exercise comparisons.

Adherence to exercise interventions was measured in three quarters of included RCTs ( $n = 12$ ; Ansai et al., 2016; Bårdstu et al., 2020; Bechshoft et al., 2017; Bonnefoy et al., 2003; Campbell et al., 1997; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Hamdorf & Penhall, 1999; Hvid et al., 2016; Kalapotharakos et al., 2010; Rosie & Taylor, 2007). Adherence was most commonly measured by attendance at supervised sessions (10 RCTs; Ansai et al., 2016; Bårdstu et al., 2020; Bechshoft et al., 2017; Bonnefoy et al., 2003; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Hamdorf & Penhall, 1999; Hvid et al., 2016; Kalapotharakos et al., 2010). Adherence to exercise sessions varied widely from 51% to 92% (mean attendance) of the scheduled sessions (Table 1).

## Outcomes

Follow-up periods ranged from 2 weeks to 24 months (median follow-up: 6 months). Physical function was reported in 14 RCTs, of which 13 used performance-based measures (Ansai et al., 2016; Bårdstu et al., 2020; Bechshoft et al., 2017; Bonnefoy et al., 2003; Campbell et al., 1997; Cancela Carral et al., 2017; de Bruin & Murer, 2007; Gine-Garriga et al., 2010; Hartshorn et al., 2002; Hvid et al., 2016; Kalapotharakos et al., 2010; Kim et al., 2019; Rosie & Taylor, 2007; Siemonsma et al., 2018) and five used self-reported measures (Bechshoft et al., 2017; Campbell et al., 1997; Cancela Carral et al., 2017; Luukinen et al., 2006; Rosie & Taylor, 2007). Health-related quality of life was reported in one RCT only (Gine-Garriga, Guerra, & Unnithan, 2013). Psychosocial outcomes

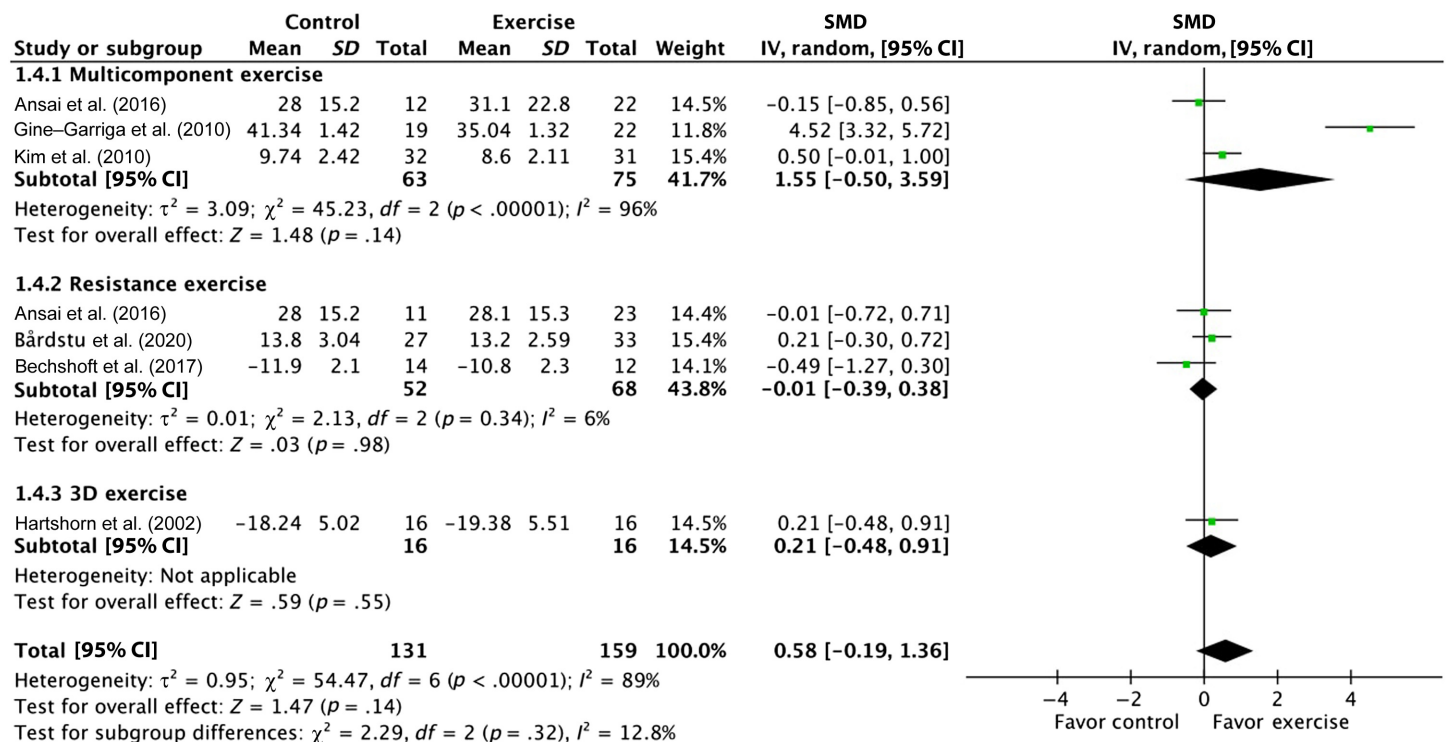
(depressive symptoms, morale, or mood) were reported in four RCTs (Ansai & Rebelatto, 2015; Hamdorf & Penhall, 1999; Hartshorn et al., 2002; Luukinen et al., 2006). Falls were reported in four RCTs (Ansai et al., 2016; Campbell et al., 1997; Luukinen et al., 2006; Rosie & Taylor, 2007).

### Therapeutic exercise compared to nonexercise comparators.

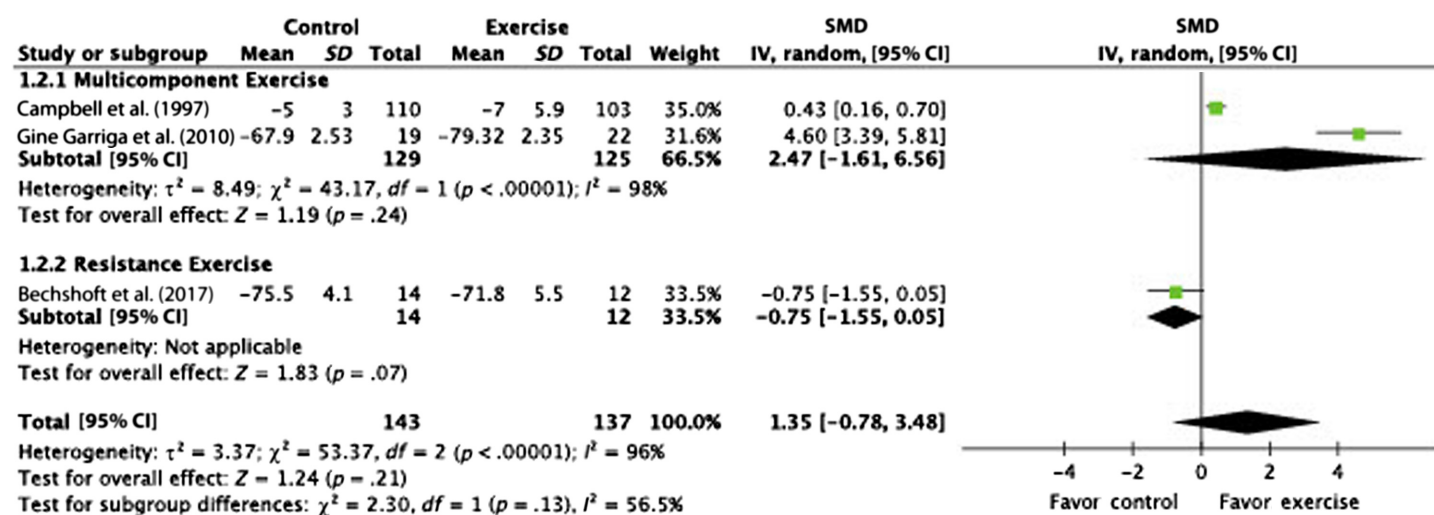
**Performance-based physical function:** Nine RCTs reported on performance-based physical function, most commonly using the Timed Up and Go Test ( $n = 7$ ). There was no evidence of a difference in effect between therapeutic exercise and nonexercise comparators on performance-based physical function at short-medium term follow-up (SMD: 0.58, 95% CI [-0.19, 1.36];  $I^2$ : 89%; 6 RCTs; 290 participants; very low-quality evidence; Figure 2). This finding was consistent across exercise types. No data were reported for long-term follow-up.

**Self-reported physical function:** Five RCTs reported self-reported measures of physical function. There was no evidence of a difference in effect between therapeutic exercise and nonexercise comparators on self-reported physical function at short-medium term follow-up (SMD: 1.35, 95% CI [-0.78, 3.48];  $I^2$ : 96%; three RCTs; 280 participants; very low-quality evidence; Figure 3). None of the RCTs included in meta-analysis used the same outcome measure, contributing to the observed high level of heterogeneity. This finding was consistent across exercise types. No data were reported for long-term follow-up.

**Health-related quality of life:** The single RCT that reported health-related quality of life (Gine-Garriga et al., 2013;  $n = 51$ ) found statistically significant benefits of a multicomponent exercise intervention compared to a nonexercise control on the physical function subscale of the 12-Item Short Form Survey, and the physical and mental composite scores at both short- and medium-term follow-ups (Table 2).



**Figure 2** — Forest plot for the effect of therapeutic exercise compared to a nonexercise comparator group on performance-based physical function. SMD = standardized mean difference; CI = confidence interval.



**Figure 3** — Forest plot for the effect of therapeutic exercise compared to a nonexercise comparator group on self-reported physical function. SMD = standardized mean difference; CI = confidence interval.

**Psychosocial outcomes:** Data from the four RCTs reporting psychosocial outcomes (Ansai & Rebelatto, 2015; Hamdorf & Penhall, 1999; Hartshorn et al., 2002; Luukinen et al., 2006) were not able to be combined in meta-analysis due to significant heterogeneity in the constructs being measured and reporting of results (Table 2).

Findings varied between the included RCTs. Ansai et al. (Ansai & Rebelatto, 2015) reported no differences in scores on the Geriatric Depression Scale at medium-term follow-up between participants who undertook 16 weeks of multicomponent training or no intervention control. Luukinen et al. (Luukinen et al., 2006) also used the Geriatric Depression Scale but used it to categorize if participants were depressed or not, and found no difference in the proportion of participants classified as being depressed at long-term follow-up between those who undertook a 16-month multicomponent program and those who continued routine care. Hamdorf et al. (Hamdorf & Penhall, 1999) found that participants who completed a 6-month aerobic exercise program had significant improvements on the Philadelphia Geriatric Morale Scale compared to a non-exercise control at medium-term follow-up (Table 2).

**Falls:** Data from the three RCTs reporting falls could not be combined in meta-analysis due to heterogeneity in follow-up periods, measurement and reporting (Table 2; Ansai et al., 2016; Campbell et al., 1999, 1997; Luukinen et al., 2007). Two RCTs (Ansai et al., 2016; Luukinen et al., 2007) reported no difference in the risk of falls at short or medium follow-up between those who completed multicomponent exercise programs or no intervention control. In contrast, Campbell et al. (1999, 1997) reported significant reductions in the relative hazard for a first fall with injury (heart rate: 0.61, 95% CI [0.39, 0.97]) and for all falls following a multicomponent home exercise program at long-term follow-up (heart rate: 0.69, 95% CI [0.49, 0.97]).

**Adverse events and mortality:** Adverse event data were reported in five RCTs ( $n = 232$ ; Ansai et al., 2016; Bårdstu et al., 2020; Bechshoft et al., 2017; Hamdorf & Penhall, 1999; Kim et al., 2019). Bechshoft et al. (2017) was the only included RCT that reported a serious adverse event: a compression fracture of epicondylus medialis femoris triggered by the resistance training. Four of the five RCTs (Bårdstu et al., 2020; Bechshoft et al., 2017;

Hamdorf & Penhall, 1999; Kim et al., 2019) reported no nonserious adverse events in either intervention or control groups. Ansai et al. (2016) reported six nonserious adverse events (mild muscle pain, mild hematoma, and dizziness) in the exercise intervention and none in the nonexercise control group.

Mortality data were reported in eight RCTs ( $n = 1,323$ ; Bårdstu et al., 2020; Bonnefoy et al., 2003; Campbell et al., 1997; Gine-Garriga et al., 2010; Hamdorf & Penhall, 1999; Kim et al., 2019; Luukinen et al., 2006; Siemonsma et al., 2018). Two RCTs reported no deaths in either intervention or control groups (Hamdorf & Penhall, 1999; Kim et al., 2019). Meta-analysis showed evidence that therapeutic exercise may reduce the risk of mortality during follow-up compared to nonexercise comparators (RR: 0.47, 95% CI [0.32, 0.70];  $I^2 = 0.00\%$ ; six RCTs; 1,222 participants; low-quality evidence; Figure 4).

**Between-exercise comparisons.** *Performance-based physical function:* Four RCTs ( $n = 163$ ; four comparisons) compared aerobic, functional, or multicomponent exercise interventions to resistance exercise comparators. There was no evidence of a difference in effect between therapeutic exercise and resistance exercise comparators on performance-based physical function at short-medium term follow-up (SMD: 0.03, 95% CI [-0.28, 0.33]; four RCTs; 163 participants;  $I^2 = 0.00\%$ ; low-quality evidence; Figure 5). This finding was consistent across exercise types. No data were reported for long-term follow-up.

*Self-reported physical function:* Two RCTs ( $n = 193$ ; two comparisons) compared aerobic or functional exercise interventions to resistance exercise comparators. There was no evidence of a difference in effect between therapeutic exercise and resistance exercise comparators on self-reported physical function at short-medium term follow-up (SMD: 0.14, 95% CI [-0.55, 0.27]; two RCTs; 193 participants;  $I^2 = 0.00\%$ ; low-quality evidence; Figure 6). This finding was consistent across exercise types. No data were reported for long-term follow-up.

*Health-related quality of life:* None of the included RCTs comparing therapeutic exercise to other exercise reported health-related quality of life outcomes.

*Psychosocial outcomes:* One RCT reported psychosocial outcomes (Ansai & Rebelatto, 2015). Ansai and Rebelatto (2015)

**Table 2 Health-Related Quality of Life, Psychosocial Outcomes, and Falls Results of Included Studies**

Study	Outcome used	Follow-up time point	Number of participants analyzed	Intervention Mean (SD)	Control Mean (SD)	Mean difference between groups [95% CI]
Health-related quality of life Gine-Garriga et al. (2010)	SF-12 physical function subscale	Short (12 weeks)	Intervention: 22 Control: 19	40.8 (6.95)	28.24 (6.49)	12.56 [8.44, 16.68] $p < .001$
		Medium (36 weeks)	Intervention: 18 Control: 7	40.3 (7.37)	29.47 (5.93)	10.83 [5.27, 16.39] $p = .001$
	SF-12 physical composite score	Short (12 weeks)	Intervention: 22 Control: 19	35.59 (4.41)	29.80 (3.74)	5.79 [3.30, 8.28] $p < 0.001$
		Medium (36 weeks)	Intervention: 18 Control: 7	36.52 (4.47)	29.26 (3.05)	7.26 [4.20, 10.32] $p = .002$
	SF-12 mental composite score	Short (12 weeks)	Intervention: 22 Control: 19	38.37 (7.14)	31.14 (8.56)	7.23 [2.36, 12.10] $p = .001$
		Medium (36 weeks)	Intervention: 18 Control: 7	33.94 (6.18)	30.53 (7.41)	3.41 [-2.78, 9.60] $p = .048$
Psychosocial outcomes Ansai and Rebelatto (2015)	GDS	Medium (16 weeks)	Multicomponent: 22 Resistance: 23 Control: 23	Multicomponent: 4.0 (2.8) Resistance: 3.2 (2.5)	3.7 (2.6)	Multicomponent vs. control 0.30 [-1.28, 1.88] Multicomponent vs. resistance 0.80 [-0.75, 2.35] Resistance vs. control -0.50 [-1.97, 0.97]
			Multicomponent: 22 Resistance: 23 Control: 23	Multicomponent: 3.7 (2.5) Resistance: 3.7 (2.8)	3.9 (2.7)	Multicomponent vs. control -0.20 [-1.72, 1.32] Multicomponent vs. resistance 0.00 [-1.55, 1.55] Resistance vs. control -0.70 [-2.29, 0.89] $p = .002$
	Philadelphia Geriatric Morale Scale	Medium (6 months)	Intervention: 18 Control: 20	9.9 (SEM: 0.38)	7.8 (SEM: 0.57)	—
	POMS	Short (2 weeks)	Intervention: 16 Control: 16	20.39 (10.57)	20.45 (11.01)	
	GDS	Long (28 months)	Intervention: 144 Control: 150	Classified as depressed ( $n$ [%]): 45 (33)	Classified as depressed ( $n$ [%]): 41 (28)	
	Frequency of feeling lonely	Long (28 months)	Intervention: 144 Control: 150	Classified as lonely ( $n$ [%]): 50 (36)	Classified as lonely ( $n$ [%]): 57 (39)	NR

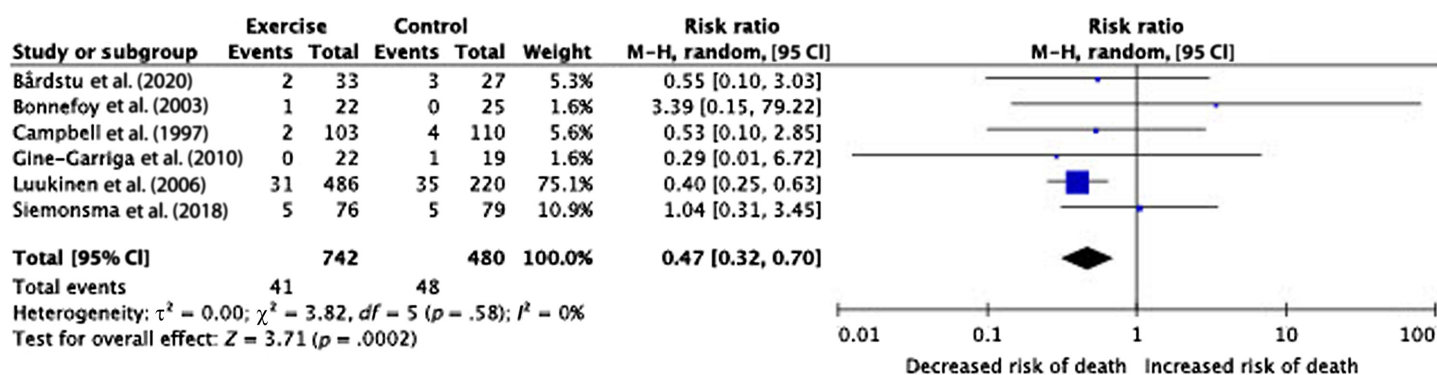
(continued)

Table 2 (continued)

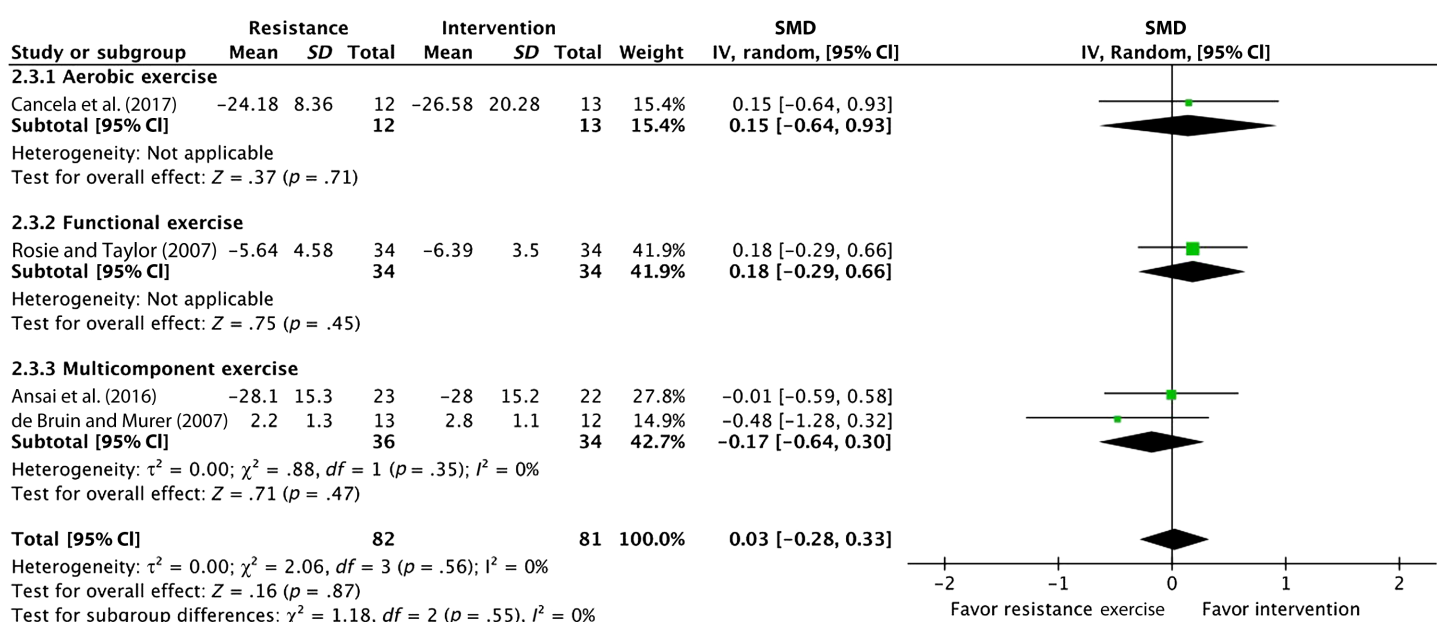
Study	Outcome used	Follow-up time point	Number of participants analyzed	Intervention group number of falls n (%)	Control group number of falls n (%)	Risk ratio [95% CI]
Falls results						
Ansai et al. (2016)	Monthly phone calls	Medium (16 weeks)	Multicomponent: 23 Resistance: 23 Control: 23	Multicomponent: 4 (18.2) Resistance: 8 (34.8)	8 (34.8)	Multicomponent vs. control 0.50 [0.17, 1.43] Multicomponent vs. resistance 0.50 [0.17, 1.43] Resistance vs. control 1.00 [0.45, 2.21]
		Medium (22 weeks)	Multicomponent: 23 Resistance: 23 Control: 23	Multicomponent: 2 (9.1) Resistance: 1 (4.3)	5 (25)	Multicomponent vs. control 0.40 [0.09, 1.86] Multicomponent vs. resistance 2.00 [0.19, 20.55] Resistance vs. control 0.20 [0.03, 1.58]
Campbell et al. (1997)	Falls calendar and telephone call	Long (12 months)	Intervention: 103 Control: 110	88 Rate (SD): 0.87 (1.29) per year	152 Rate (SD): 1.34 (1.94) per year	Reported as hazard ratios First fall: 0.81 [0.56, 1.16] First four falls: 0.68 [0.52, 0.90] Fall resulting in moderate or severe injury: 0.61 [0.39, 0.97]
Campbell et al. (1999)	Falls calendar and telephone call	Long (24 months)	Intervention: 41 Control: 62	138 Rate (SD): 0.83 per year	220 Rate (SD): 1.19 per year	Reported as hazard ratios All falls: 0.69 [0.49, 0.97] Fall resulting in moderate or severe injury: 0.63 [0.42, 0.95]
Luukinen et al. (2007)	Phone calls and medical record confirmation	Bimonthly for 3 years	Intervention: 217 Control: 220	Rate [95% CI]: 1.15 [1.03, 1.29] per person year	Rate [95% CI]: 1.23 [1.10, 1.37] per person year	First 4 falls: 0.88 [0.74, 1.04] All falls: 0.93 [0.80, 1.09]
Rosie and Taylor (2007)	Diary and telephone call	Short (6 weeks)	Functional: 34 Resistance: 34	2	3	0.67 [0.12, 3.74]

Note. SF-12 = 12-Item Short Form Survey; GDS = Geriatric Depression Scale; POMS = profile of mood states; CI = confidence interval.

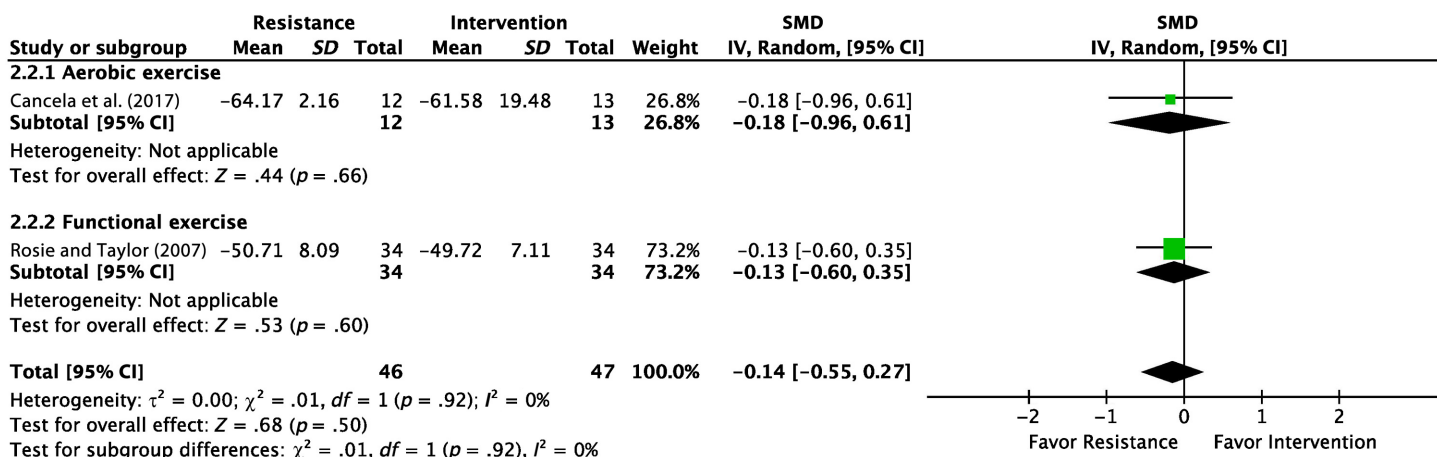




**Figure 4** — Forest plot for the effect of therapeutic exercise compared to a nonexercise comparator group on mortality. CI = confidence interval.



**Figure 5** — Forest plot for the effect of therapeutic exercise compared to a resistance exercise comparator group on performance-based physical function. SMD = standardized mean difference; CI = confidence interval.



**Figure 6** — Forest plot for the effect of therapeutic exercise compared to a resistance exercise comparator group on self-reported physical function. SMD = standardized mean difference; CI = confidence interval.



found no significant differences in scores on the Geriatric Depression Scale at medium-term follow-up between participants who undertook a multicomponent training or resistance training (mean difference: 0.00, 95% CI [-1.55, 1.55]); Table 2).

**Falls:** One RCT reported falls. Rosie and Taylor (2007) reported no difference in the number of falls short-term between participants who undertook a functional training program and those who completed resistance exercises (RR: 0.67, 95% CI [0.12, 3.74]; Table 2).

**Adverse events and mortality:** Adverse event data were reported in two RCTs ( $n = 113$ ; Ansai et al., 2016; Rosie & Taylor, 2007). No serious adverse events were reported. All nonserious adverse events reported were mild muscle pain. Meta-analysis showed no difference in risk of nonserious adverse events between therapeutic exercise and resistance exercise comparators (RR: 0.79, 95% CI [0.40, 1.57];  $I^2 = 0.00\%$ ; two RCTs; 113 participants; moderate-quality evidence; Figure 7). Mortality data were reported

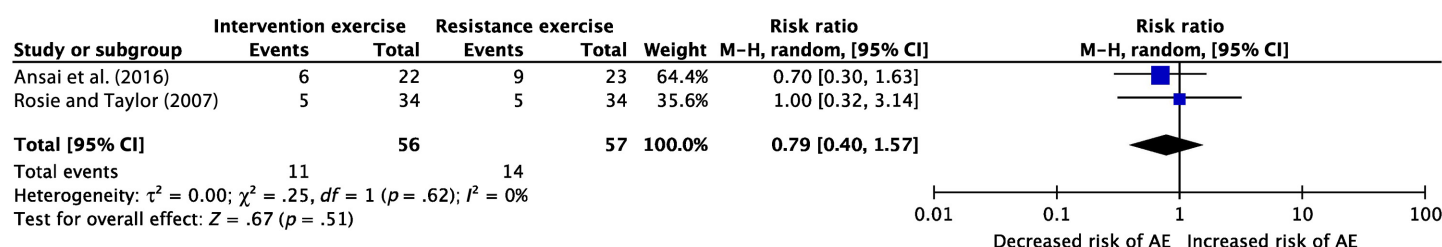
in one RCT that compared multicomponent exercise to resistance exercise (de Bruin & Murer, 2007). No deaths were reported in either exercise group.

## Overall Evaluation of the Quality of Evidence

The overall quality of the evidence assessed using GRADE, including reasons for downgrading, is summarized in Table 3. The certainty of evidence for all outcomes was downgraded for risk of bias and inconsistency, resulting in low confidence in the effect estimate.

## Discussion

We identified 16 RCTs including 1,660 participants evaluating therapeutic exercise interventions among community-dwelling



**Figure 7** — Forest plot for the effect of therapeutic exercise compared to a resistance exercise comparator group on nonserious adverse events. AE = adverse event; CI = confidence interval.

**Table 3 Summary of Findings**

Outcome	Risk with nonexposed comparator	Risk with therapeutic exercise	Anticipated absolute effects (SMD) [95% CI]	Relative effect [95% CI]	Number of participants (studies)	Certainty of the evidence (GRADE)
Therapeutic exercise vs. nonexercise comparator						
Performance-based physical function	—	—	0.58 higher [0.19 lower to 1.36 higher]	—	290 (six RCTs)	●●●○ Very low <sup>a,b,c</sup>
Self-reported physical function	—	—	1.35 higher [0.78 lower to 3.48 higher]	—	280 (three RCTs)	●●●○ Very low <sup>a,b,c</sup>
Mortality	55 per 1,000	10 per 1,000	—	0.47 [0.32, 0.70]	1,222 (six RCTs)	●●●○ Low <sup>a,c</sup>
Therapeutic exercise vs. resistance exercise comparator						
Performance-based physical function	—	—	0.03 higher [0.28 lower to 0.33 higher]	—	163 (four RCTs)	●●●○ Low <sup>a,b</sup>
Self-reported physical function	—	—	0.14 lower [0.55 lower to 0.27 higher]	—	193 (two RCTs)	●●●○ Low <sup>a,b</sup>
Nonserious adverse events	246 per 1,000	196 per 1,000	—	0.79 [0.40, 1.57]	113 (two RCTs)	●●●○ Moderate <sup>a</sup>

**Note.** GRADE working group grades of evidence. High certainty: We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect. Very low certainty: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect. CI = confidence interval; GRADE = Grading of Recommendations, Assessment, Development and Evaluation; RCTs = randomized controlled trials; SF-12 = 12-Item Short Form Survey; SMD = standardized mean difference.

<sup>a</sup>Downgraded one level for risk of bias (more than one trial at high or unclear risk of bias). <sup>b</sup>Downgraded one level for inconsistency (considerable statistical heterogeneity in these outcomes that could not be explained by prespecified sensitivity and subgroup analyses). <sup>c</sup>Downgraded one level for imprecision (relatively broad overall CI).

adults aged 80 years or older. The trials included a range of exercise interventions, most commonly multicomponent or resistance exercise, with most trials using a nonexercise control group.

Due to significant heterogeneity across RCTs, we were only able to conduct a limited meta-analysis. We found no evidence of a difference in effect between therapeutic exercise and nonexercise comparators on performance-based or self-reported physical function at short-medium term follow-up. There was some evidence that therapeutic exercise reduced the risk of mortality during follow-up compared to nonexercise controls. Individual trials reported significant benefits of therapeutic exercise on health-related quality of life, psychosocial outcomes and falls, however these findings have to be considered uncertain as many of these measures were secondary outcomes.

## Results in Context

To our knowledge, no systematic reviews have been published evaluating the effects of therapeutic exercise on physical function and psychosocial outcomes among community-dwelling adults aged 80 years or older. The small number of RCTs that were included in our review was unsurprising, given that people aged 80 and over have been deemed “the great forgotten” in clinical studies (Izquierdo et al., 2020).

Our finding of a lack of evidence for the effectiveness of therapeutic exercise interventions on physical function, health-related quality of life, and psychosocial outcomes highlights the gap between the known physiological benefits of exercise and observed clinical benefits in published RCTs. Among the nine included RCTs that reported exercise intensity, both heart rate and resistance intensities were most commonly moderate. We cannot be certain that the therapeutic exercise was sufficient to cause physiological changes needed to improve strength and function. Previous reviews have reported that exercise is commonly prescribed to older adults at an insufficient dose to improve strength and function (Singh, 2002; Steib, Schoene, & Pfeifer, 2010). Future research examining the effects of different dosages of therapeutic exercise is required.

The variation in results of the included RCTs is consistent with a recent systematic review of meta-analyses on exercise interventions in older adults aged  $\geq 65$  or mean age  $\geq 70$  (Di Lorito et al., 2021). Di Lorito et al. also found variable results and no significant effects for performance-based physical function (Di Lorito et al., 2021). We noted that health-related quality of life was only measured in one included RCT and psychosocial outcomes were only reported in 25% of included RCTs in our review, a finding also noted by Di Lorito et al. (2021).

Our findings suggest that therapeutic exercise may reduce mortality for community-dwelling adults aged 80 years or older. This is consistent with a recent systematic review and meta-analysis of the safety and effectiveness of long-term exercise interventions among older adults ( $n = 28,532$ , mean age 74.2 years), which found that long-term exercise was associated with reduced risk of serious adverse events (including mortality; García-Hermoso et al., 2020). Generally high rates of completion and exercise adherence across the included studies also suggest that exercise interventions are acceptable and feasible for people aged 80 years or older. This supports the similar finding from an earlier systematic review of exercise interventions for frail older people living in the community (Clegg, Barber, Young, Forster, & Iliffe, 2012).

## Implications for Clinical Practice

High-quality evidence demonstrating the effectiveness of therapeutic exercise for adults aged 80 years or older is lacking. While future RCTs are undertaken to improve our confidence in the effects of exercise for this population, given that exercise seems safe, it should be recommended in clinical practice. Which type of therapeutic exercise is most beneficial remains unclear, so exercise selection should be based on an individual's deficits, goals, and preferences.

## Implications for Future Research

This systematic review highlights the need for high-quality RCTs evaluating the effects of well-designed therapeutic exercise interventions among community-dwelling adults aged 80 years or older. Identifying how best to facilitate participation of this population in such trials is an important issue. The Innovations in Clinical Trial Design and Delivery for the Under-served project seeks to address this issue, and the resources produced from this work should guide the design of future RCTs (Witham et al., 2020). Future RCTs should assess health-related quality of life, psychosocial outcomes, and cost-effectiveness analyses in addition to physical function, and should include long-term follow-up. Care should also be given to the content of the comparator intervention (Freedland et al., 2019). Variation in comparator interventions may in part explain the inconsistent effects observed in our review.

Given that adults aged 80 years or older are a highly heterogeneous population interventions should be individualized and tailored to the goals and preferences of each participant. Exercise intervention design should be based on behavior change theory, and should include behavior change techniques such as goal setting, motivational strategies, and booster sessions, to optimize exercise uptake and adherence long term (Nicolson et al., 2017).

## Strengths and Limitations

The strengths of this review include the comprehensive search, explicit eligibility criteria, duplicate assessment of eligibility, risk of bias assessment, and use of the GRADE approach to rate the overall quality of evidence.

The limitations of our review relate largely to the underlying evidence. First, the heterogeneous nature of the included RCTs reflected the large inconsistency in the effect estimates of the meta-analyses. Second, the sample sizes of included RCTs were generally small, and many included outcomes were secondary. Finally, the quality of evidence for the outcomes of interest was low to very low, suggesting that the true effect may differ from the findings reported in this review.

## Conclusions

Results of this systematic review and meta-analysis suggest that evidence for the effectiveness of therapeutic exercise interventions on physical function, health-related quality of life and psychosocial outcomes in community-dwelling adults aged 80 years or older is limited, and of low quality. Therapeutic exercise may reduce mortality in this population. Well-designed RCTs should be a research priority.

## Acknowledgments

We would like to sincerely thank Liz Callow, Outreach Librarian at the Bodleian Health Care Libraries, University of Oxford for her assistance with building the search strategy for this review.

PN conceived the study idea, gained funding, was involved in screening, data extraction and data analysis and led the writing of the review. SH, EW and SL contributed to study conception, provided methodological guidance, contributed to data analysis and writing of the review. VD was involved in screening, data extraction and data analysis, and contributed to writing of the review. SL acts as the guarantor of the review. All authors gave their final approval of the version to be published.

PN is funded by a Versus Arthritis Foundation Fellowship (Grant Reference Number 22428). SL and EW receive funding from the Collaboration for Leadership in Applied Health Research and Care Oxford at Oxford Health National Health Service (NHS) Foundation Trust and are supported by National Institute for Health Research (NIHR) Musculoskeletal Biomedical Research Centre, Oxford. The funders had no role in the design; collection, analysis or reporting of this review. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

## References

- Ansai, J.H., Aurichio, T.R., Goncalves, R., & Rebelatto, J.R. (2016). Effects of two physical exercise protocols on physical performance related to falls in the oldest old: A randomized controlled trial. *Geriatrics and Gerontology International*, 16(4), 492–499. PubMed ID: 25868484 doi:10.1111/ggi.12497
- Ansai, J.H., & Rebelatto, J.R. (2015). Effect of two physical exercise protocols on cognition and depressive symptoms in oldest-old people: A randomized controlled trial. *Geriatrics and Gerontology International*, 15(9), 1127–1134. PubMed ID: 25407380 doi:10.1111/ggi.12411
- Bårdstu, H.B., Andersen, V., Fimland, M.S., Aasdahl, L., Raastad, T., Cumming, K.T., & Sæterbakken, A.H. (2020). Effectiveness of a resistance training program on physical function, muscle strength, and body composition in community-dwelling older adults receiving home care: A cluster-randomized controlled trial. *European Review of Aging and Physical Activity*, 17(1), 11. doi:10.1186/s11556-020-00243-9
- Bechshoft, R.L., Malmgaard-Clausen, N.M., Gliese, B., Beyer, N., Mackey, A.L., Andersen, J.L., ... Holm, L. (2017). Improved skeletal muscle mass and strength after heavy strength training in very old individuals. *Experimental Gerontology*, 92, 96–105. PubMed ID: 28363433 doi:10.1016/j.exger.2017.03.014
- Bonnefoy, M., Cornu, C., Normand, S., Boutitie, F., Bugnard, F., Rahmani, A., ... Laville, M. (2003). The effects of exercise and protein-energy supplements on body composition and muscle function in frail elderly individuals: A long-term controlled randomised study. *British Journal of Nutrition*, 89(5), 731–738. doi:10.1079/BJN2003836
- Campbell, A.J., Robertson, M.C., Gardner, M.M., Norton, R.N., & Buchner, D.M. (1999). Falls prevention over 2 years: A randomized controlled trial in women 80 years and older. *Age and Ageing*, 28(6), 513–518. PubMed ID: 10604501 doi:10.1093/ageing/28.6.513
- Campbell, A.J., Robertson, M.C., Gardner, M.M., Norton, R.N., Tilyard, M.W., & Buchner, D.M. (1997). Randomised controlled trial of a general practice programme of home based exercise to prevent falls in elderly women. *BMJ*, 315(7115), 1065–1069. PubMed ID: 9366737 doi:10.1136/bmj.315.7115.1065
- Cancela Carral, J.M., Pallin, E., Orbegozo, A., & Ayan Perez, C. (2017). Effects of three different chair-based exercise programs on people older than 80 years. *Rejuvenation Research*, 20(5), 411–419. PubMed ID: 28482740 doi:10.1089/rej.2017.1924
- Clegg, A.P., Barber, S.E., Young, J.B., Forster, A., & Iliffe, S.J. (2012). Do home-based exercise interventions improve outcomes for frail older people? Findings from a systematic review. *Reviews in Clinical Gerontology*, 22(1), 68–78. PubMed ID: 27226701 doi:10.1017/S0959259811000165
- The Cochrane Collaboration. (2019). Cochrane handbook for systematic reviews of interventions version 6.0 (Updated July 2019). [www.handbook.cochrane.org](http://www.handbook.cochrane.org)
- Collerton, J., Jagger, C., Yadegarfar, M.E., Davies, K., Parker, S.G., Robinson, L., & Kirkwood, T.B.L. (2016). Deconstructing complex multimorbidity in the very old: Findings from the Newcastle 85+ Study. *Biomed Research International*, 2016, 8745670. PubMed ID: 26885519 doi:10.1155/2016/8745670
- de Bruin, E.D., & Murer, K. (2007). Effect of additional functional exercises on balance in elderly people. *Clinical Rehabilitation*, 21(2), 112–121. PubMed ID: 17264105 doi:10.1177/0269215506070144
- Di Lorito, C., Long, A., Byrne, A., Harwood, R.H., Gladman, J.R.F., Schneider, S., ... van der Wardt, V. (2021). Exercise interventions for older adults: A systematic review of meta-analyses. *Journal of Sport and Health Science*, 10(1), 29–47. PubMed ID: 32525097 doi:10.1016/j.jshs.2020.06.003
- England, P.H. (2018). Health Profile for England. Retrieved from <https://www.gov.uk/government/publications/health-profile-for-england-2018>
- Fransen, M., McConnell, S., Harmer, A.R., Van der Esch, M., Simic, M., & Bennell, K.L. (2015). Exercise for osteoarthritis of the knee. *Cochrane Database Systematic Reviews*, 9(1), CD004376. doi:10.1002/14651858.CD004376.pub3
- Fransen, M., McConnell, S., Hernandez-Molina, G., & Reichenbach, S. (2014). Exercise for osteoarthritis of the hip. *Cochrane Database Systematic Reviews*, 22(4), CD007912.
- Freedland, K.E., King, A.C., Ambrosius, W.T., Mayo-Wilson, E., Mohr, D.C., Czajkowski, S.M., ... Riley, W.T. (2019). The selection of comparators for randomized controlled trials of health-related behavioral interventions: Recommendations of an NIH expert panel. *Journal of Clinical Epidemiology*, 110, 74–81. PubMed ID: 30826377 doi:10.1016/j.jclinepi.2019.02.011
- García-Hermoso, A., Ramirez-Vélez, R., Sáez de Asteasu, M.L., Martínez-Velilla, N., Zambom-Ferraresi, F., Valenzuela, P.L., ... Izquierdo, M. (2020). Safety and effectiveness of long-term exercise interventions in older adults: A systematic review and meta-analysis of randomized controlled trials. *Sports Medicine*. doi:10.1007/s40279-020-01259-y
- Gine-Garriga, M., Guerra, M., Pages, E., Manini, T.M., Jimenez, R., & Unnithan, V.B. (2010). The effect of functional circuit training on physical frailty in frail older adults: A randomized controlled trial. *Journal of Aging and Physical Activity*, 18(4), 401–424. PubMed ID: 20956842
- Gine-Garriga, M., Guerra, M., & Unnithan, V.B. (2013). The effect of functional circuit training on self-reported fear of falling and health status in a group of physically frail older individuals: A randomized controlled trial. *Aging Clinical and Experimental Research*, 25(3), 329–336. PubMed ID: 23740589 doi:10.1007/s40520-013-0048-3
- Goodpaster, B.H., Park, S.W., Harris, T.B., Kritchevsky, S.B., Nevitt, M., Schwartz, A.V., ... Newman, A.B. (2006). The loss of skeletal muscle strength, mass, and quality in older adults: The health, aging and body composition study. *The Journals of Gerontology. Series A*,



- Biological Sciences & Medical Sciences*, 61(10), 1059–1064. PubMed ID: 17077199 doi:10.1093/gerona/61.10.1059
- Hamdorf, P.A., & Penhall, R.K. (1999). Walking with its training effects on the fitness and activity patterns of 79–91 year old females. *Australian and New Zealand Journal of Medicine*, 29(1), 22–28. PubMed ID: 10200809 doi:10.1111/j.1445-5994.1999.tb01584.x
- Hartshorn, K., Delage, J., Field, T., & Olds, L. (2002). Senior citizens benefit from movement therapy. *Journal of Bodywork & Movement Therapies*, 6(1), 55–58. doi:10.1054/jbmt.2001.0229
- Heyn, P.C., Johnson, K.E., & Kramer, A.F. (2008). Endurance and strength training outcomes on cognitively impaired and cognitively intact older adults: A meta-analysis. *Journal of Nutrition, Health and Aging*, 12(6), 401–409. doi:10.1007/BF02982674
- Higgins, J.P.T., Altman, D.G., Gtzsche, P.C., Jni, P., Moher, D., Oxman, A.D., . . . Sterne, J.A.C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343, d5928. doi:10.1136/bmj.d5928
- Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V.A. (2020). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Retrieved from [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook)
- Hughes, V.A., Frontera, W.R., Wood, M., Evans, W.J., Dallal, G.E., Roubenoff, R., & Fiatarone Singh, M.A. (2001). Longitudinal muscle strength changes in older adults: Influence of muscle mass, physical activity, and health. *The Journals of Gerontology. Series A, Biological Sciences & Medical Sciences*, 56(5), B209–B217. PubMed ID: 11320101 doi:10.1093/gerona/56.5.B209
- Hvid, L.G., Strotmeyer, E.S., Skjodt, M., Magnussen, L.V., Andersen, M., & Caserotti, P. (2016). Voluntary muscle activation improves with power training and is associated with changes in gait speed in mobility-limited older adults—A randomized controlled trial. *Experimental Gerontology*, 80, 51–56. PubMed ID: 27090485 doi:10.1016/j.exger.2016.03.018
- Izquierdo, M., Morley, J.E., & Lucia, A. (2020). Exercise in people over 85. *BMJ*, 368, m402. PubMed ID: 32024635 doi:10.1136/bmj.m402
- James Lind Alliance. (2018). Multiple conditions in later life top 10 priorities. Retrieved from <http://www.jla.nihr.ac.uk/priority-setting-partnerships/health-with-multiple-conditions-in-old-age/top-10-priorities.htm>
- Kalapotharakos, V.I., Diamantopoulos, K., & Tokmakidis, S.P. (2010). Effects of resistance training and detraining on muscle strength and functional performance of older adults aged 80 to 88 years. *Aging Clinical and Experimental Research*, 22(2), 134–140. PubMed ID: 20440099 doi:10.1007/BF03324786
- Kim, H., Won, C.W., Kim, M., Kojima, N., Fujino, K., Osuka, Y., . . . Suzuki, T. (2019). The effects of exercise and milk-fat globule membrane (MFGM) on walking parameters in community-dwelling elderly Japanese women with declines in walking ability: A randomized placebo controlled trial. *Archives of Gerontology & Geriatrics*, 83, 106–113. PubMed ID: 30999125 doi:10.1016/j.archger.2019.03.029
- Lafortune, L., Béland, F., Bergman, H., & Ankri, J. (2009). Health state profiles and service utilization in community-living elderly. *Medical Care*, 47(3), 286–294. PubMed ID: 19165121 doi:10.1097/MLR.0b013e3181894293
- Lamb, S.E., Becker, C., Gillespie, L.D., Smith, J.L., Finnegan, S., Potter, R., . . . Taxonomy, I. (2011). Reporting of complex interventions in clinical trials: Development of a taxonomy to classify and describe fall-prevention interventions. *Trials*, 12(1), 125. PubMed ID: 21586143 doi:10.1186/1745-6215-12-125
- Luukinen, H., Lehtola, S., Jokelainen, J., Vaananen-Sainio, R., Lotvonen, S., & Koistinen, P. (2006). Prevention of disability by exercise among the elderly: A population-based, randomized, controlled trial. *Scandinavian Journal of Primary Health Care*, 24(4), 199–205. PubMed ID: 17118858 doi:10.1080/02813430600958476
- Luukinen, H., Lehtola, S., Jokelainen, J., Väänänen-Sainio, R., Lotvonen, S., & Koistinen, P. (2007). Pragmatic exercise-oriented prevention of falls among the elderly: A population-based, randomized, controlled trial. *Preventive Medicine*, 44(3), 265–271. PubMed ID: 17174387 doi:10.1016/j.ypmed.2006.09.011
- Manini, T.M., & Pahor, M. (2009). Physical activity and maintaining physical function in older adults. *British Journal of Sports Medicine*, 43(1), 28–31. PubMed ID: 18927164 doi:10.1136/bjsm.2008.053736
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269. PubMed ID: 19622511 doi:10.7326/0003-4819-151-4-200908180-00135
- Nicolson, P.J.A., Bennell, K.L., Dobson, F.L., Van Ginckel, A., Holden, M.A., & Hinman, R.S. (2017). Interventions to increase adherence to therapeutic exercise in older adults with low back pain and/or hip/knee osteoarthritis: A systematic review and meta-analysis. *British Journal of Sports Medicine*, 51(10), 791–799. PubMed ID: 28087567 doi:10.1136/bjsports-2016-096458
- Office for National Statistics. (2012). Changes in the older resident care home population between 2001 and 2011. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/changesintheolderresidentcarehomepopulationbetween2001and2011/2014-08-01>
- Office for National Statistics. (2018). National population projections: 2018-based. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2018based#changing-age-structure>
- Rosie, J., & Taylor, D. (2007). Sit-to-stand as home exercise for mobility-limited adults over 80 years of age—GrandStand System may keep you standing? *Age and Ageing*, 36(5), 555–562. PubMed ID: 17646216 doi:10.1093/ageing/afm093
- Santoni, G., Angleman, S., Welmer, A.-K., Mangialasche, F., Marengoni, A., & Fratiglioni, L. (2015). Age-related variation in health status after age 60. *PLoS One*, 10(3), e0120077. doi:10.1371/journal.pone.0120077
- Siemonsma, P.C., Blom, J.W., Hofstetter, H., van Hespren, A.T.H., Gussekloo, J., Drewes, Y.M., & van Meeteren, N.L.U. (2018). The effectiveness of functional task exercise and physical therapy as prevention of functional decline in community dwelling older people with complex health problems. *BMC Geriatrics*, 18(1), 164. PubMed ID: 30016948 doi:10.1186/s12877-018-0859-3
- Silva, R.B., Aldoradin-Cabeza, H., Eslick, G.D., Phu, S., & Duque, G. (2017). The effect of physical exercise on frail older persons: A systematic review. *The Journal of Frailty & Aging*, 6(2), 91–96. PubMed ID: 28555710 doi:10.14283/jfa.2017.7
- Singh, M.A. (2002). Exercise comes of age: Rationale and recommendations for a geriatric exercise prescription. *The Journals of Gerontology. Series A, Biological Sciences & Medical Sciences*, 57(5), M262–M282. PubMed ID: 11983720 doi:10.1093/gerona/57.5.M262
- Smith, T.O., Collier, T.S., Smith, B., & Mansfield, M. (2019). Who seeks physiotherapy or exercise treatment for hip and knee osteoarthritis? A cross-sectional analysis of the English longitudinal study of age cohort. *International Journal of Rheumatic Diseases*, 22(5), 897–904. PubMed ID: 30729705 doi:10.1111/1756-185X.13480
- Steib, S., Schoene, D., & Pfeifer, K. (2010). Dose-response relationship of resistance training in older adults: A meta-analysis. *Medicine &*

- Science in Sports & Exercise*, 42(5), 902–914. PubMed ID: [19996996](#) doi:[10.1249/MSS.0b013e3181c34465](#)
- Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., . . . Higgins, J.P.T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, 366, 14898. PubMed ID: [31462531](#) doi:[10.1136/bmj.l4898](#)
- U.S. Food and Drug Administration. (2020). What is a serious adverse event? Retrieved from <https://www.fda.gov/safety/reporting-serious-problems-fda/what-serious-adverse-event>
- Witham, M.D., Anderson, E., Carroll, C., Dark, P.M., Down, K., Hall, A.S., . . . Rochester, L. (2020). Developing a roadmap to improve trial delivery for under-served groups: Results from a UK multi-stakeholder process. *Trials*, 21(1), 694. PubMed ID: [32738919](#) doi:[10.1186/s13063-020-04613-7](#)
- World Health Organization. (2010). Global recommendations on physical activity for health. Retrieved from <https://www.who.int/dietphysicalactivity/global-PA-recs-2010.pdf>