**Handgrip Strength and Health Outcomes: Umbrella Review of Systematic Reviews with Meta-Analyses of Observational Studies**

**Short Title: Handgrip Strength and Health Outcomes**

Pinar Soysal\*1, Christopher Hurst\*2, Jacopo Demurtas3, Joseph Firth4, Reuben Howden5, Lin Yang6, Mark A Tully7, Ai Koyanagi8, 9, Petre Cristian Ilie10, Guillermo F. López-Sánchez11, Lukas Schwingshackl12, Nicola Veronese± 13, Lee Smith± 14

\*Joint first author

±Joint senior and corresponding authors

1. Department of Geriatric Medicine, Bezmialem Vakif University, Istanbul, Turkey

2. Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK

3. Primary Care Department, Azienda Usl Toscana Sud Est, Grosseto, Italy

4. Division of Psychology and Mental Health, University of Manchester, Manchester, UK

5. College of Health & Human Services, UNC Charlotte, USA.

6. Department of Cancer Epidemiology and Prevention Research, Alberta Health Services, Holy Cross Centre, Canada.

7. School of Health Sciences, Institute of Mental Health Sciences, Ulster University, Shore Road, Newtownabbey, Co Antrim, Northern Ireland.

8. Research and Development Unit, Parc Sanitari Sant Joan de Déu, CIBERSAM, Barcelona, Spain

9. ICREA, Pg. Lluis Companys 23, Barcelona, Spain

10. The Queen Elizabeth Hospital King's Lynn NHS Foundation Trust, UK.

11. Faculty of Sport Sciences, University of Murcia, Murcia, Spain.

12. Institute for Evidence in Medicine, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany.

13. Neuroscience Institute, Aging Branch, National Research Council, Padua, Italy.

14. The Cambridge Centre for Sport and Exercise Sciences, Anglia Ruskin University, Cambridge, UK.

Corresponding authors: Dr Lee Smith; The Cambridge Centre for Sport and Exercise Sciences, Anglia Ruskin University, Cambridge, UK; [lee.smith@anglia.ac.uk](mailto:lee.smith@anglia.ac.uk); Dr Nicola Veronese; Neuroscience Institute, Aging Branch, National Research Council, Padua, Italy; [ilmannato@gmail.com](mailto:ilmannato@gmail.com)

**Conflicts of Interest:** None. **Funding:** None.

**ABSTRACT**

**Purpose:** The aim of the present study was to assess both the credibility and strength of evidence arising from systematic reviews with meta-analyses of observational studies on handgrip strength and health outcomes.

**Methods:** An umbrella review of systematic reviews with meta-analyses of observational studies was conducted. We assessed meta-analyses of observational studies based on random-effect summary effect sizes and their p values, 95% prediction intervals, heterogeneity, small-study effects and excess significance, grading the evidence from convincing (class I) to weak (IV).

**Results:** From 504 articles returned, 8 systematic reviews were included for a total of 11 outcomes. Overall, 9/11 of the outcomes reported nominally significant summary results (p<0**.**05), with four associations surviving the application of the more stringent p-value (*P* < 10−6). No outcome presented convincing evidence. Three associations showed class II evidence (i.e. highly suggestive): higher handgrip values at the baseline were associated with a minor reduction in mortality risk in the general population (n=34 studies; sample size: 1,855,817; risk ratio, RR=0.72; 95% confidence intervals, CI: 0.67-0.78) and cardiovascular death risk in mixed populations (n=15 studies; RR=0.84; 95%CI: 0.78-0.91) and incidence of disability (n=7 studies; RR=0.76; 95%CI: 0.66-0.87).

**Conclusion:** The present results show that handgrip strength is a useful indicator for general health status and specifically early all-cause and cardiovascular mortality, as well as disability. Future research is now required to fully understand mechanisms linking handgrip strength scores to these health outcomes to further inform intervention strategies.

**Key words:** Handgrip strength, Health outcomes, Meta-analysis, Umbrella Review.

**Introduction**

A decline in physical function is a natural phenomenon that is associated with aging.1 Such a decline is a public health concern as it has been shown to be associated with increased risk of falls,2 health care use,3 level of dependency,4 and premature mortality.5 Indeed, for many independent older adults’ everyday tasks such as climbing stairs require functioning close to maximal capacity, meaning that further decline could increase their risk of becoming dependent on a carer.6 One widely employed measure of physical functioning is handgrip strength. The handgrip strength test is commonly used to evaluate the integrated performance of the muscles by determining the maximal grip force that can be produced in one muscular contraction, further serving as a marker for general muscle strength.7 Handgrip strength is a valid measure of physical function and has been widely employed in observational research and clinical settings.8-11 Importantly, one study found that dynamometer determined handgrip strength could be a useful instrument in geriatric practice to identify the ‘oldest old’ patients (i.e. those aged over 75 years) at risk of disability.12

In recent years there has been an exponential increase in the literature investigating associations between handgrip strength and health outcomes (e.g. depression,13 cognitive function14, suicidal ideation15, mobility limitations,16 falls,17 cardiovascular disease,18 diabetes,19 renal outcomes,20 osteoporotic factors,21 multi-morbidity,22 and mortality23) and consequently an increase in systematic reviews with meta-analyses. However, to date, most systematic reviews have focused on a single disease end point and there has not been a systematic evaluation of the relationships between handgrip strength and diverse physical and mental health outcomes. Moreover, the strength and reliability of the literature is unclear.

In order to address the breadth of the literature of physiological measurements and outcomes, an increasing emphasis has been placed on “umbrella reviews” (i.e., the syntheses of existing systematic reviews with meta-analyses, to capture the breadth of outcomes associated with a given exposure).

Given this, the aim of the present study was to carry out an umbrella review of existing systematic reviews with meta-analyses of handgrip strength and all health outcomes in order to systematically assess the quality and strength of the evidence across all health outcomes and to identify those studies with the strongest evidence.

**Methods**

This umbrella review was registered in PROSPERO

(<https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=158547>).

*Data sources and searches*

We conducted an umbrella review,24 searching several databases (MEDLINE, Scopus, Embase) from inception until 20th November 2019 with: “(Meta-Analysis[ptyp] OR metaanaly\*[tiab] OR meta-analy\*[tiab] OR Systematic review [ptyp] OR “systematic review” [tiab]) AND (handgrip [tiab])”. In addition, we hand-searched the reference lists of eligible articles.

*Study selection*

In this umbrella review, we included: systematic reviews and/or meta-analyses of observational studies that investigate the relationship between handgrip strength and any health outcome. Specific inclusion criteria include: 1) Meta-analyses or systematic reviews containing sufficient data for a meta-analysis (as defined by the authors), measuring handgrip strength, and ascertaining health outcomes using self-report (e.g. depression questionnaire), observed (e.g. clinical diagnosis) or objective (e.g. biomarkers, mortality) criteria; 2) case control, or cohort studies (retrospective and prospective cohorts); 3) meta-analyses of cohort studies that investigated the association between handgrip strength with any health-related outcome (e.g. cardiovascular disease, cancer, death, obesity/overweight, mental illness, diabetes and metabolic diseases). Studies had to report these outcomes as odds ratio (OR), relative risk (RR), hazard ratio (HR) or continuous data. Two authors (PS, CH) independently performed title and abstract screening in couples. Disagreements were resolved through consensus with another independent author (LS).

*Data extraction*

Four independent investigators (PS, LS, CH, NV) extracted in pairs the following information for each article: first author name; year of publication; journal; the number of included studies and the total number of people included in the review; the inclusion criteria for studied population; the measure(s) by which handgrip strength was captured; how handgrip strength was categorized; the effect size(s) used in the review; sub-grouping used in the meta-analysis; study design (case-control, retrospective, prospective); number of cases and controls for each study; and health outcome.

We then extracted the study-specific estimated relative risk for health outcomes (RR, OR, HR, SMDs, i.e. standardized mean difference), along with the 95% confidence intervals (CI), and the number of cases for each study by subjects and controls. If two works were available for the same association, we included the largest in terms of studies.

*Risk of Bias Assessment*

Two authors (PS, CH) independently rated the methodological quality of the included systematic reviews using the AMSTAR 2 tool.25 The AMSTAR 2 tool ranks the quality of a meta-analysis from critically low to high according to 16 predefined items. The review is ranked as high quality if it has no or one non-critical weakness - the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest. The review is ranked as moderate quality if it has more than one non-critical weakness - the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review. The review is ranked as low quality if it has one critical flaw with or without non-critical weaknesses - the review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest. Finally, the review is ranked as critically low quality if it has more than one critical flaw with or without non-critical weaknesses - the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies.26 For further reading relating to the AMSTAR 2 and what constitutes a critical flaw or a critical weakness etc. we refer the reader to the following reference.26

*Statistical Analysis*

For each meta-analysis, we estimated the summary effect size and its 95% CI through random-effects models.27 We also estimate the prediction interval (PIs) and its 95% CI, which further accounts for between-study effects, and estimates the certainty of the association if a new study addresses that same association.28 Between-study association was estimated with the I² metric; values >50% is indicative of high heterogeneity, while values above 75% suggests very high heterogeneity.29

In addition, we calculated the evidence of small-study effects (i.e., whether small studies would have inflated effect sizes compared to larger ones). To this end, we used the regression asymmetry test developed by Egger and co-workers.30 A P value < 0.10 with more conservative effects in larger studies than in random-effects meta-analysis will be considered as indicative of small-study effects.21 Finally, we applied the Ioannidis’ excess of significance test to evaluate whether there was an excess of studies reporting statistically significant results.31

*Grading the Evidence*

We used the credibility assessment criteria, which are based on established tools for observational evidence as summarized previously.24,32-35 We classified evidence from meta-analyses of observational studies with nominally statistically significant summary results (P<0.05) into four categories (class I, II, III, and IV). We considered convincing associations with a statistical significance of P<10-6, including more than 1,000 cases (or more than 20,000 participants for continuous outcomes), have the largest component study reporting a significant result (P<0.05), have a 95% prediction interval that excluded the null, did not have large heterogeneity (I² <50%), and showed no evidence of small study effects (P>0.10) and of excess significance bias (P>0.10). Highly suggestive (class II) evidence was assigned to associations that reported a significance of P<0.001, include more than 1,000 cases (or more than 20,000 participants for continuous outcomes), and had the largest component study reporting a statistically significant result (<0.05). Suggestive (class III) evidence was assigned to associations that reported a significance of P<0.01 with more than 1,000 cases (or more than 20,000 participants for continuous outcomes). Weak (class IV) evidence was assigned to the remaining significant associations with P<0.05.

Due to the inherent limitations of case-control studies to examine temporal associations, we planned to provide the classification of evidence, for class I and II, based on the following order: a) meta-analyses of prospective studies b) meta-analyses of prospective and retrospective case-control studies. However, no outcome had these characteristics.

**Results**

*Literature Review*

We identified 20 potentially eligible papers of these, eight publications were selected as eligible with eleven different outcomes being included in this umbrella review.

*Meta-analyses of Observational Studies*

The median number of studies of meta-analyses including observational studies for each outcome was eight (range 4-34), the median number of participants was 23,064 (range 275 to 1,855,817), and the median number of cases was 1,823 (**Table 1**).

The majority of the meta-analyses included studies on general population or in adults older than 50 years, followed by patients with cardiovascular disease. Overall, nine of the 11 outcomes reported nominally significant summary results (p<0**.**05), with four associations surviving to the application of the more stringent p-value (*P* < 10−6) (**Table 1**). Heterogeneity among studies was high in nine of the 11 of the outcomes included, with seven having an I2 > 75%. Only two associations presented 95% PIs excluding the null value. Evidence for excess statistical significance was present in five of 41 outcomes and small-study effects were also seen in three of 11 of the outcomes. Bias was present in three of the outcomes included. The largest study, in terms of participants for each outcome, was statistically significant in all the associations, except one.

Based on the above criteria, no outcome presented convincing evidence. However, three associations showed class II evidence (i.e. highly suggestive): higher handgrip values at the baseline, in fact were associated with a minor reduction in mortality risk in the general population (n=34 studies; sample size: 1,855,817; RR=0.72; 95%CI: 0.67-0.78) and cardiovascular death in mixed populations (e.g. diabetes, general, other conditions) (n=15 studies; RR=0.84; 95%CI: 0.78-0.91) and, finally, incidence of disability (n=7 studies; RR=0.76; 95%CI: 0.66-0.87) (**Table 1**). The other outcomes were ranked as convincing (association between higher handgrip values and chair rise performance over time) or weak (five outcomes), with only two associations not statistically significant (i.e. the association between handgrip strength and incident hip fracture or cancer mortality) (**Table 1**).

*Quality Assessment*

Utilizing the AMSTAR 2 tool a total of four included meta-analyses scored “critically low” and four “low” (Supplementary Table 1). Notably most studies did not include a list of excluded studies (n=8) or report the source of funding for the included studies (n=7). Moreover, it should be noted one study did not include a systematic review.

**Discussion**

In this umbrella review of 8 meta-analyses and 11 health outcomes investigating associations between handgrip strength and all health outcomes, a total of 3 outcomes (lower all-cause mortality, lower cardiovascular mortality, and lower risk of disability) were found to have highly suggestive evidence. One outcome (chair rise performance over time) was found to have convincing evidence. Five outcomes were found to have a weak evidence (walking speed, inability to balance, hospital admissions, cardiac death, and mortality in those with chronic kidney disease). Importantly, two associations were found to be non-significant (incident hip fracture and cancer mortality). Taken together these findings suggest that handgrip strength is a useful indicator for general health status and specifically early all-cause and cardiovascular mortality, as well as disability and leg power (chair rise performance).

Several mechanisms may explain the relationship between handgrip strength and early mortality. First, early life factors, such as participation in sufficient levels of physical activity influence handgrip strength36 and childhood levels of physical activity and handgrip strength have been shown to track into adulthood.37,38 Importantly, maintaining adequate levels of physical activity and function over the entire life course likely yields the greatest benefit to health owing to the reduction of any prolonged exposure from unhealthy behaviours. Next, strength is related to muscle mass and muscle mass is as a protein reserve during cases of trauma.39 Finally, other genetic contributions may be at play that result in muscle dystrophy and early mortality.40

When considering the relationship between handgrip strength and disability and leg power, this may be explained by sarcopenia (a progressive reduction in muscle strength and mass absolute and relative to body size commonly occurring with aging).41 Sarcopenia is associated with a decline in physical function and an increase in disability.8 Next, the handgrip strength test is not just a pure measure of strength and those with joint disorders, that will likely increase risk of disability and lower leg power, may perform worse when carrying out this task.8

Umbrella reviews provide top-tier evidence and important insights, but there are a number of limitations that need to be considered. The meta-analyses contained studies that differed in their designs, populations, and other characteristics. However, we applied an I2 <50% as 1 of the criteria for Class I evidence (convincing), in order to assign the best evidence grade only to robust associations. Next, meta-analyses have inherent limitations42: their findings depend on which estimates are selected from each primary study and how they are applied in the meta-analysis. Finally, all included meta-analyses in the present review scored low or critically low when appraised using the AMSTAR 2 tool suggesting that future meta-analyses in this area require more accurate reporting of methods and results as well as incorporating more robust discussions around findings.

**Conclusions**

In conclusion, the present results show that handgrip strength is a useful indicator for general health status and specifically early all-cause and cardiovascular mortality, as well as disability and leg power (chair rise performance). Future research is now required to fully understand mechanisms linking handgrip strength scores to these health outcomes to further inform intervention strategies.

**Conflicts of Interest:** None.

**Funding:** None.

**References**

1. Runge M, Rittweger J, Russo CR, Schiessl H, Felsenberg D. Is muscle power output a key factor in the age-related decline in physical performance? A comparison of muscle cross section, chair-rising test and jumping power. *Clin Physiol Funct Imaging*. 2004;24(6):335–40.

2. Singh DKA, Shahar S, Vanoh D, Kamaruzzaman SB, Tan MP. Diabetes, arthritis, urinary incontinence, poor self‐rated health, higher body mass index and lower handgrip strength are associated with falls among community‐dwelling middle‐aged and older adults: Pooled analyses from two cross‐sectional Malaysian datasets. *Geriatr Gerontol Int*. 2019;19(8):798–803.

3. Cheng Y, Goodin AJ, Pahor M, Manini T, Brown JD. Healthcare utilization and physical functioning in older adults in the United States. *J Am Geriatr Soc*. 2020;68(2):266–71.

4. Meskers CGM, Reijnierse EM, Numans ST, Kruizinga RC, Pierik VD, et al. Association of handgrip strength and muscle mass with dependency in (instrumental) activities of daily living in hospitalized older adults -The EMPOWER Study. *J Nutr Health Aging*. 2019;23(3):232–8.

5. Eekhoff EMW, van Schoor NM, Biedermann JS, Oosterwerff MM, De Jongh R, et al. Relative importance of four functional measures as predictors of 15-year mortality in the older Dutch population. *BMC Geriatr*. 2019;19(1):92.

6. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. *J Aging Phys Act*. 1999;7(2):129–61.

7. Leong DP, Teo KK, Rangarajan S, Kutty VR, Lanas F, et al. Reference ranges of handgrip strength from 125,462 healthy adults in 21 countries: a prospective urban rural epidemiologic (PURE) study. *J Cachexia Sarcopenia Muscle*. 2016;7(5):535–46.

8. Giampaoli S. Hand-grip strength predicts incident disability in non-disabled older men. *Age Ageing*. 1999;28(3):283–88.

9. Rantanen T. Midlife hand grip strength as a predictor of old age disability. *JAMA*. 1999;281(6):558.

10. Rantanen T, Avlund K, Suominen H, Schroll M, Frändin K, Pertti E. Muscle strength as a predictor of onset of ADL dependence in people aged 75 years. *Aging Clin Exp Res*. 2002;14(3 Suppl):10–5.

11. Onder G, Penninx BWJH, Ferrucci L, Fried LP, Guralnik JM, Pahor M. Measures of physical performance and risk for progressive and catastrophic disability: Results from the Women’s Health and Aging Study. *J Gerontol A Biol Sci Med Sci*. 2005;60(1):74–9.

12. Taekema DG, Gussekloo J, Maier AB, Westendorp RGJ, de Craen AJM. Handgrip strength as a predictor of functional, psychological and social health. A prospective population-based study among the oldest old. *Age Ageing*. 2010;39(3):331–7.

13. Smith L, White S, Stubbs B, Hu L, Veronese N, et al. Depressive symptoms, handgrip strength, and weight status in US older adults. *J Affect Disord*. 2018;238:305–10.

14. Yang L, Koyanagi A, Smith L, Hu L, Colditz GA, et al. Hand grip strength and cognitive function among elderly cancer survivors. *PloS one* 2018;13(6):e0197909.

15. Cao C, Liu Q, Yang L, Zheng X, Lan P, et al. Handgrip strength is associated with suicidal thoughts in men: Cross‐sectional analyses from NHANES. *Scand J Med Sci Spor.* 2020;30(1):92-9.

16. Sallinen J, Stenholm S, Rantanen T, Heliövaara M, Sainio P, et al. Hand‐grip strength cut points to screen older persons at risk for mobility limitation. *J Am Geriatr Soc.* 2010;58(9):1721-6.

17. Hoda W, Samia AR, Ahmed M. Handgrip strength and falls in community-dwelling Egyptian seniors. *Adv Aging Res*. 2013;2(4):37618.

18. Celis-Morales CA, Welsh P, Lyall DM, Steell L, Petermann F, et al. Associations of grip strength with cardiovascular, respiratory, and cancer outcomes and all cause mortality: prospective cohort study of half a million UK Biobank participants. *BMJ*. 2018;361:k1651.

19. Karvonen-Gutierrez CA, Peng Q, Peterson M, Duchowny K, Nan B, et al. Low grip strength predicts incident diabetes among mid-life women: the Michigan Study of Women’s Health Across the Nation. *Age Ageing* 2018;47(5):685-91.

20. Chang YT, Wu HL, Guo HR, Cheng YY, Tseng CC, et al. Handgrip strength is an independent predictor of renal outcomes in patients with chronic kidney diseases. *Nephrol Dial Transpl* 2011;26(11):3588-95.

21. Cheung CL, Tan KC, Bow CH, Soong CS, Loong CH, et al. Low handgrip strength is a predictor of osteoporotic fractures: cross-sectional and prospective evidence from the Hong Kong Osteoporosis Study. *Age* 2012;34(5):1239-48.

22. Cheung CL, Nguyen US, Au E, Tan KC, Kung AW. Association of handgrip strength with chronic diseases and multimorbidity. *Age* 2013;35(3):929-41.

23. Kim GR, Sun J, Han M, Park S, Nam CM. Impact of handgrip strength on cardiovascular, cancer and all-cause mortality in the Korean longitudinal study of ageing. *BMJ Open*. 2019;9(5):e027019.

24. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ Can Med Assoc J*. 2009;181(8):488–93.

25. Wells G, Shea B, O’Connell D, Robertson J, Peterson J, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. *Ottawa: Ottawa Hospital Research Institute*. 2015.

26. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.

27. Lau J. Quantitative synthesis in systematic reviews. *Ann Intern Med*. 1997;127(9):820-826.

28. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):137–59.

29. Ioannidis JPA, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*. 2007;335(7626):914–16.

30. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–34.

31. Ioannidis JP. Clarifications on the application and interpretation of the test for excess significance and its extensions. *J. Math. Psychol.* 2013;57(5):184-7.

32. Veronese N, Solmi M, Caruso MG, Giannelli G, Osella A, et al. Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses. *Am J Clin Nutr*. 2018;107(3):436–44.

33. Veronese N, Demurtas J, Celotto S, Caruso MG, Maggi S, et al. Is chocolate consumption associated with health outcomes? An umbrella review of systematic reviews and meta-analyses. *Clin Nutr*. 2019;38(3):1101–08.

34. Smith L, Luchini C, Demurtas J, Soysal, P, Stubbs B, et al. Telomere length and health outcomes: An umbrella review of systematic reviews and meta-analyses of observational studies. *Ageing Res Rev*. 2019;51:1–10.

35. Veronese N, Demurtas J, Pesolillo G, Celotto S, Barnini T, et al. Magnesium and health outcomes: an umbrella review of systematic reviews and meta-analyses of observational and intervention studies. *Eur J Nutr*. 2020;59(1):263–72.

36. Kim S-H, Lim B-O, An K-O. Association of physical activity and handgrip strength among Korean elderly. *Asian J Kinesiol*. 2019;21(4):16–21.

37. Trudeau F, Shephard RJ, Arsenault F, Laurencelle L. Tracking of physical fitness from childhood to adulthood. *Can J Appl Physiol*. 2003;28(2):257–71.

38. Smith L, Gardner B, Aggio D, Hamer M. Association between participation in outdoor play and sport at 10years old with physical activity in adulthood. *Prev Med*. 2015;74:31–5.

39. Buckner SL, Dankel SJ, Bell ZW, Abe T, Loenneke JP. The association of handgrip strength and mortality: what does it tell us and what can we do with it? *Rejuvenation Res*. 2019;22(3):230–4.

40. Salzberg DC, Mann JR, McDermott S. Differences in race and ethnicity in muscular dystrophy mortality rates for males under 40 years of age, 2006–2015. *Neuroepidemiology*. 2018;50(3–4):201–6.

41. Santilli V, Bernetti A, Mangone M, Paoloni M. Clinical definition of sarcopenia. *Clin Cases Miner Bone Metab*. 2014;11(3):177–80.

42. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94(3):485–514.

**Table 1. Health outcomes and evidence class reported in included meta-analyses of observational studies.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Author of MA** | **Population** | **Outcome** | **Study design** | **Number of**  **studies** | **Cases** | **Sample size** | **Effect size** | **Mean ES (95%CI)** | **P** | **I2** | **Small study effects** | **Excess significance bias** | **Largest study significant** | **95% PI** | **Class of evidence** |
| García-Hermoso A  2018 | General population | All-Cause Mortality | Cohort | 34 | 57854 | 1855817 | RR | 0.72  (0.67-0.78) | 2.04E-18 | 83.5 | yes | no | yes | 0.52-1 | **II** |
| Chainani V,  2016 | Mixed | CVD mortality | Cohort/clinical trials | 15 | 2183 | 29105 | RR | 0.84  (0.78-0.91) | 0.00001 | 84.3 | no | no | yes | 0.67-1.07 | **II** |
| Vermeulen,  2011 | Mixed | Disability | Cohort | 7 | 1136 | 5201 | RR | 0.76  (0.66-0.87) | 0.00009 | 89.9 | yes | yes | yes | 0.5-1.16 | **II** |
| Hardy R,  2013 | Adults aged 50 years or older | Chair rise performance | Cohort | 8 | NA | 10098 | beta | 0.93  (0.65-1.21) | 6.20E-11 | 91 | no | NA | yes | -0.02-1.88 | **III** |
| Hardy R,  2013 | Adults aged 50 years or older | Walking speed | Cohort | 8 | NA | 7261 | beta | 0.89  (0.61-1.17) | 5.37E-10 | 88.4 | no | NA | yes | -0.03-1.82 | **IV** |
| Hardy R,  2013 | Adults aged 50 years or older | Inability to balance | Cohort | 8 | NA | 11318 | OR | 0.94  (0.92-0.98) | 1.58E-09 | 76.2 | no | NA | yes | 0.88-1 | **IV** |
| Pavasini R, 2018 | Patients with cardiac disease | Cardiac death | Cohort/clinical trials | 6 | 3000 | 23435 | OR | 0.83  (0.74-0.94) | 0.01 | 52.1 | no | yes | yes | 0.59-1.17 | **IV** |
| Pavasini R, 2018 | Patients with cardiac disease | Hospital admission for HF | Cohort/clinical trials | 4 | 125 | 23064 | OR | 0.88  (0.82-0.95) | 0.01 | 14.3 | no | no | yes | 0.71-1.10 | **IV** |
| Hwang SH, 2019 | Patients With CKD Undergoing Dialysis | Mortality | Cohort | 10 | 589 | 2775 | RR | 0.92  (0.87-0.98) | 0.02 | 70.3 | yes | no | yes | 0.85-1.19 | **IV** |
| Denk K, 2018 | Adults aged 50 years or older | Hip fracture | Case-control | 12 | 1462 | 28579 | RR | 1.32  (0.97-1.79) | 0.08 | 90.8 | no | yes | yes | 0.5-3.47 | **NS** |
| García-Hermoso A, 2018 | Healthy youth and adults | Cancer mortality | Cohort | 10 | 8887 | 1297163 | RR | 0.97  (0.92-1.07) | 0.28 | 18.9 | no | no | no | 0.88-1.07 | **NS** |

**Supplementary Table: AMSTAR 2 evaluation of included studies**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Year** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **AMSTAR 2 Rating** |
| Hardy | 2013 | No | No | No | No | No | No | No | Yes | No | No | Yes | No | No | Yes | No | Yes | Critically Low |
| Chainani | 2016 | No | No | No | Partial Yes | Yes | No | No | Yes | No | No | No | No | Yes | No | Yes | Yes | Critically Low |
| Vermeulen | 2011 | Yes | No | Yes | Partial Yes | Yes | Yes | No | Yes | No | No | No | No | No | No | No | Yes | Critically Low |
| Garcia-Hermoso | 2018 | Yes | Yes | Yes | Partial Yes | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | No | Yes | No | Critically Low |
| Pavasini | 2018 | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Low |
| Hwang | 2019 | Yes | No | No | Partial Yes | Yes | No | No | No | Yes | No | No | No | No | Yes | Yes | Yes | Low |
| Garcia-Hermoso | 2018 | Yes | Yes | No | Partial Yes | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | Yes | Low |
| Denk | 2018 | Yes | Yes | Yes | Partial Yes | Yes | No | No | Yes | Yes | Yes | No | No | No | Yes | No | Yes | Low |

A Measurement Tool to Assess systematic Reviews (AMSTAR) 2 questions; *1. Did the research questions and inclusion criteria for the review include the components of PICO?, 2.* Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?, 3. Did the review authors explain their selection of the study designs for inclusion in the review?, *4. Did the review authors use a comprehensive literature search strategy?, 5. Did the review authors perform study selection in duplicate?, 6. Did the review authors perform data extraction in duplicate?, 7. Did the review authors provide a list of excluded studies and justify the exclusions?, 8. Did the review authors describe the included studies in adequate detail?, 9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?, 10. Did the review authors report on the sources of funding for the studies included in the review?, 11. If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results?, 12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?, 13. Did the review authors account for RoB in primary studies when interpreting/discussing the results of the review?, 14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?, 15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?, 16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?*