Supplementary Table Summary of included studies on the association between PA and QoL, described by study design

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Effect size	Main findings
Intervention studies								
King[24]	19 points (70/high)	U.S./not reported	Randomized; 103 subjects (65% women); over 65 years of age	Duration: 12 months Description: Two weekly sessions of oriented exercises (1 hour) and at least two workout sessions at home IG_1: Exercises focused on muscular strength and endurance IG_2: Exercises focused on muscular strength and flexibility	HRQoL and domains of MOS (physical and psychological well-being)	None	- IG_1: 0.54 (energy/ fatigue). - IG_2: 0.94 (bodily pain) and 0.58 (emotional well-being).	- The indicators of HRQoL physical and psychological well-being remained stable after 12 months of intervention in both groups. - The only indicators that improved after the intervention were energy/fatigue in IG_1 and emotional well-being and bodily pain in IG_2.
Varejão[25]	18 points (63/moderate)	Brazil/not reported	Not randomized; 69 healthy women; between 60 and 76 years of age	Duration: 24 weeks Description: Three weekly sessions (1 hour). The training program of both groups consisted of 41 exercises, with two sets of 10 seconds spent on each movement IG_1: Exercises focused on flexibility IG_2: Exercises focused on stretching	General QoL and domains of WHOQoL-100	None	- IG_1: 0.36 (physical); 0.02 (psychological); 0.19 (personal relationships); 0.2 (environmental; 0.11 (spiritual). - IG_2: 0.14 (physical); 0.27 (psychological); 0.18 (personal relationships); 0.18 (environmental); 0.25 (spiritual).	- There were no significant increases in WHOQoL-100 domain scores, with small differences between pretest and posttest of flexing in the QoL domains physical (2.27%), psychological (0.05%), personal relationships (1.78%), environmental (1.8%), and spiritual (1.9%); and of stretching in the QoL domains physical (1.2%), psychological (2.1%), personal relationships (2.0%), environmental (1.2%), and spiritual (3.8%).
Fox[26]	21 points (78/high)	England/not reported	Not randomized; 176 subjects (56% women); mean age 75 years	Duration: 12 months Description: Three weekly sessions (60-90 minutes) IG: The program included strength, flexibility, and aerobic training. Two weekly sessions were conducted in group meetings, while another weekly session took place at home. PA was measured with an accelerometer. CG	General QoL and WHOQoL-Bref domains	None	- The study did not provide sufficient data to calculate the effect size.	- The higher the energy expenditure and minutes spent in moderate PA per day, the higher the scores of general QoL and on the physical ($r = 0.22$, $p < 0.05$), environmental ($r = 0.25$, $p < 0.01$), social ($r = 0.08$, $p > 0.05$), and psychological domains ($r = 0.21$, $p < 0.01$). Light PA was not significantly correlated with general QoL scores or the WHOQoL-100 domains (baseline data analysis) physical ($r = 0.11$), environmental ($r = 0.05$), social ($r = 0.04$), or psychological ($r = 0.04$). Follow-up measures (12 months) demonstrated no statistically significant changes in well-being or QoL scores. Physical domain scores remained stable for men in the IG, but decreased in the CG ($p < 0.01$).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Effect size	Main findings
Vreede[27]	19 points (70/high)	Netherlands/ not reported	Randomized; 98 women; over 70 years of age	Duration: 12 weeks (with 9-month follow-up after intervention) Description: Three weekly sessions (1 hour) IG_1: Functional exercises IG_2: Endurance exercises CG	HRQoL and SF-36 domains	None	- IG_2 (3 months): 0.7 (physical function). - IG_1 (3 and 9 months): 0.7 (physical component); 0.09 (mental component). - IG_1 (9 months): 1.06 (pain); 1.00 (physical component); 0.63 (physical function).	- There were no significant changes in within-group HRQoL scores after 3 and 9 months of intervention, in both IG_2 and IG_1. - Comparing the groups after 3 months, there was a greater increase in physical function domain scores in IG_2 as compared with IG_1 and CG scores. Between 3 and 9 months, summary scores for the mental and physical component domains showed greater reductions in IG_1 than in the CG. On comparison between baseline and 9 months posttest, greater reductions in pain, physical function, and summary physical component scores were observed in IG_1 as compared with the CG.
Morey[28]	19 points (70/high)	Australia/not reported	Randomized; 357 subjects (30% women) who had a recent diagnosis of cancer (prostate or breast); between 65-94 years of age	Duration: 6 months Description: Two 6-month interventions (PA and healthy-eating) were completed with elderly subjects. Actions used in the intervention were phone-based counseling and distribution of materials. The CG had no further contact. PA involvement was evaluated before and after the intervention, based on international recommendations (150+ min/week of PA)	SF-36 (only section of functional trajectory)	Age, gender, race, functional trajectory, and number of diseases	- IG: 0.47 (functional trajectory).	- After adjusting for confounding factors, changes in PA behavior after 6 months of intervention resulted in a mean change of -11.18 ($p < 0.001$) to +5.10 ($p = 0.05$), respectively, in the functional trajectory domain. The observed changes in PA behavior could be starting above the recommendation (150 min +/week) and ending below it, or starting below the recommendation and ending above it.

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Effect size	Main findings
Rocha[29]	18 points (63/moderate)	Brazil/not reported	Not randomized; 98 healthy subjects; mean age 68 years	Duration: 12 weeks Description: Two weekly sessions (50 minutes each) at moderate intensity (60-70% maximum heart rate). IG_1: Strength exercises IG_2: Muscle endurance exercises CG	General QoL and domains of WHOQoL-Old	None	- IG_1: 0.60 (autonomy); 0.77 (past, present, and future activities); 0.56 (social participation); 0.45 (intimacy); 0.56 general QoL). - IG_2: 0.38 (sensory skills); 0.41 (autonomy); 0.36 (past, present, and future activities); 0.63 (social participation); 0.62 (death and dying); 0.36 (intimacy); 1.13 (general QoL).	- IG_1 and IG_2 had significant increases ($p < 0.05$) in all domains of QoL, except sensory skills and death and dying in IG_1, which had decreases in their scores. The CG showed no changes in QoL. - IG_1 had increases in the scores of four QoL domains (autonomy, past, present, and future activities, social participation, and intimacy) and of general QoL ($p < 0.05$) when compared to the CG - IG2 had increases in the scores of six QoL domains (sensory skills, autonomy, past, present, and future activities, death and dying, social participation, and intimacy) and of general QoL ($p < 0.05$) when compared to the CG - IG_2 showed greater improvements in sensory skill, death and dying, and general QoL domain scores than the IG_1
Dec-hamps[30]	19 points (70/high)	France/2006-2008	Randomized; 160 institutionalized subjects; between 65-102 years of age (mean age 82.3 years)	Duration: 6 months (12-month follow-up after intervention) Description: IG_1: Tai chi (four weekly sessions of 30 min each) IG_2: Action-cognition program (two weekly sessions of 30-45 min each) CG	Symptoms of depression assessed by NPI	Age, gender, neuro-psychiatric diagnosis, number of comorbidities	- IG_1, IG_2, GC: 0.23 (depression after 6 months); - IG_1, IG_2: 0.29 (depression after 12 months).	- Total NPI scores increased significantly in the CG after 6 months, while remaining unchanged in IG_1 and IG_2. These differences were even higher 12 months after the intervention - Symptoms of depression decreased significantly after 6 months in IG_1, IG_2, and CG; however, these effects were maintained after 12 months only in IG_1 e IG_2 (Cohen F of 0.96 and 1.27, respectively, both with $p < 0.01$).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Effect size	Main findings
Taguchi[31]	20 points (74/high)	Japan/2004-2005	Not randomized; 65 subjects; between 74 and 96 years of age	Duration: 12 months Description: One weekly session (90 min), total of 55 sessions, with estimated intensity of the sessions according to ACSM guidelines IG: Flexibility, strength, endurance, balance, aerobic capacity, body awareness, and rhythm exercises CG	HRQoL assessed by the following instruments: IADL; PGC morale scale; GDS; TMIG-IC; FES-I	None	- The study did not provide sufficient data to calculate the effect size.	- Significant increase in physical capacity and falls efficacy scores in IG ($p < 0.05$), while remaining unchanged in the CG. - There was no significant difference between the two groups; however, the difference in falls efficacy was of borderline significance ($p = 0.06$).
Pernam-buco[32]	22 points (81/high)	Brazil/not reported	Randomized; 159 subjects (41% women); mean age 66.61 years	Duration: 16 weeks Description: Strength training for 16 weeks. IG: Active Group CG	WHOQoL-100	None	- IG: 0.21 (Dom 1); 0.22 (Dom 2); 0.23 (Dom 3); 0.19 (Dom 4); 0.15 (Dom 6); 0.20 (overall QoL).	- After 16 weeks, the CG had slight improvement in QoL levels, but the difference was not significant ($p > 0.05$). Mean Domain 1 scores decreased from pre- to posttest. - The QoL scores obtained by the IG were significantly better after 16 weeks, showing that PA promoted modifications on perceptions of life. - When analyzing the domains separately, the following domains reached significance after intervention in the IG: Dom 1 (pain and discomfort, energy, and fatigue, $p < 0.001$); Dom 2 (psychology, $p < 0.001$); Dom 3 (mobility, activities of daily living, $p < 0.006$); Dom 4 (personal relationships, social support, $p < 0.05$); Dom 6 (spirituality, $p < 0.05$); and, finally, QoL index (overall QoL, $p = 0.012$).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Effect size	Main findings
Courneya[33]	22 points (85/high)	Canada/ 2003-2007	Randomized; 320 women; aged 50-74 years	Duration: 1 year Description: The exercise group was asked to perform moderate to vigorous intensity aerobic exercise 45 min/day, 5 days/ week. IG: Aerobic exercise program CG	Domains of SF-36	Age, marital status, education, site, BMI, use of anti-depressants, and presence of comorbidities	- IG: 0.32 (general health); 0.16 (physical functioning); 0.20 (bodily pain); 0.22 (vitality).	- Compared to the CG, after adjustment variables the IG maintained significantly better physical functioning (5.2, 95%CI 2.1-8.2, $p < 0.01$), general health (4.1, 95%CI 1.6-6.6, $p < 0.01$), vitality (5.2, 95%CI 1.9-8.5, $p < 0.01$), and bodily pain (5.1, 95%CI 0.8-9.4, $p < 0.02$). - A 1-year moderate-to-vigorous aerobic exercise program prevents decline in the physical aspects of QoL in postmenopausal women
Geirsdottir[34]	22 points (85/high)	Iceland/not reported	Not randomized; 238 subjects (58.2% women); mean age 73 years	Duration: 12 weeks Description: Resistance exercise program three times/ week; 3 sets, 6-8 repetitions at 75-80% of the 1-repetition maximum	HRQL	Age, gender, habitual PA, change in lean mass and strength	- IG: 0.20 (HRQoL); 0.18 (anxiety); 0.14 (satisfaction with life); 0.18 (vitality); 0.09 (social functioning); 0.09 (general QoL).	- HRQoL improved significantly both for men and women taken together after the 12-week period of intervention (55 ± 6.3 vs. 56.2 ± 5.6) in the general QoL, social functioning ($p < 0.05$), vitality and satisfaction with life, anxiety, and HRQoL domains ($p < 0.001$). - The regression model explained 70.7% of the variance in the HRQoL endpoint, and the results showed that lower baseline PA was associated with greater improvement in HRQoL ($\beta = -0.123$, $p = 0.008$).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Longitudinal studies (cohort studies)							
McAuley[35]	20 points (74/high)	U.S./2002-2005	249 women from the community; between 59-84 years of age (mean age 68.1 years)	24-month follow-up PASE	Functional limitation evaluated with LL-FDI Mediating variable: Physical function performance is an indicator based on performance on four motor tests (normal walking speed over a 7-m pathway, ascent and descent of a flight of 15 stairs, the timed 8-Foot Up-and-Go, ability to rise from a seated position, walk a distance of 8 feet as quickly as possible). Self-efficacy assessed by The Activity-Specific Balance Confidence Scale and Self-Efficacy for Walking Scale	Age, race, income, education, and chronic health conditions	- At baseline, PA was directly associated with self-efficacy for exercise (β = 0.33) and balance (β = 0.31). These variables were associated with functional fitness (β = -0.28 and β = -0.34, respectively). Measures of self-efficacy for exercise and balance were significantly related to the physical limitations domain of QoL (β = 0.34 and β = 0.24), as well as functional fitness (β = -0.49). - PA was associated with functional fitness (β = -0.16). In addition, subjects with higher self-efficacy for exercise and functional fitness reported lower functional limitations (indicator of QoL). - In the 24-month analysis, the associations between variables were smaller but still significant (β-values between 0.08 and 0.30). Therefore, the increase of PA over time was associated with improved self-efficacy for exercise, which is associated with better functional fitness and better scores in the functional limitations domain of QoL. After controlling for confounders, the associations remained significant, without substantial changes in the consistency and direction of the association.
McAuley[36]	20 points (74/high)	U.S./2002-2005	249 women from the community; between 59-84 years of age (mean age 68.1 years)	24-month follow-up PASE	SF-12 SWLS Mediating variable: Self-efficacy assessed by The Activity-Specific Balance Confidence Scale related to daily PA practice without compromising balance and prevention of falls	Age, race, income, education and chronic health conditions	- At baseline, daily PA was associated with self-efficacy (β = 0.31). Daily PA was also related to physical health (β = 0.43) and mental health (β = 0.24). Having better physical health and mental health was associated with increased satisfaction (β = 0.38 and β = 0.37, respectively). - At 24 months, the associations between changes in variables over time were smaller but still significant (β = -0.19 to 0.38), except for the relationship between changes in physical health and changes in satisfaction. The general model explained 32% and 51% of the variance in satisfaction at baseline and 24 months respectively. - Therefore, the increase in PA over time was associated with improved self-efficacy for daily PA, which correlates with higher functional fitness. After controlling for confounders, the associations remained significant.

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Aoyagi[37]	21 points (74/high)	Japan/2002-2003	183 healthy subjects; between 65-85 years of age	12-month follow-up Pedometers (daily minutes in moderate or vigorous PA, $>$ 3 METs)	HRQoL and domains of SF-36	Age, gender, and daily step count number	- After adjustments for confounders, the mean score of the eight SF-36 domains, physical functioning, pain, vitality, and mental health were significantly higher ($p < 0.05$, $r^2 = 0.93$ and 0.87) among subjects who spent a high proportion of the PA time (over 20%) in moderate to vigorous intensity ($>$ 3 MET) as compared with those who did not participate in high-intensity PA as much.
Elavsky[38]	19 points (70/high)	U.S./not reported	174 subjects; mean age 66.7 years	Longitudinal analysis of PA and QoL in subjects who participated in the intervention, with a full battery of assessments done at 1 and 5 years after baseline PASE	General QoL assessed by SWLS Physical self-esteem assessed with PSW Exercise self-efficacy assessed with Exercise Self-Efficacy Scale Affection assessed with MUNSH	None	- After 1 year, PA had a significant and direct effect on self-efficacy for exercise ($\beta = 0.29$), physical self-esteem ($\beta = 0.38$), and affection ($\beta = 0.18$). Self-efficacy ($\beta = 0.17$) and affection ($\beta = 0.59$) had a direct effect on life satisfaction in the same period. That is, active participants, 1 year after baseline, had higher levels of physical self-esteem, self-efficacy for exercise, and affection. The last two contributed to improved global QoL during this period. - After 5 years of monitoring, changes in PA affected physical self-esteem ($\beta = 0.14$) and affection ($\beta = 0.20$), but only affection had a direct effect on global QoL (life satisfaction). The relationship between PA and global QoL was borderline significant ($\beta = 0.11$; $p = 0.052$). These results indicate that an increase in PA over time is associated with increased self-esteem and affection. Improving affection, in turn, was associated with increases in global QoL.
Cross-sectional studies							
Brown[39]	21 points (95/high)	U.S./2001	175,850 subjects, among whom 30,146 were over 65 years of age	Questionnaire developed for the BRFSS to assess PA	Questionnaire developed for the BRFSS to assess QoL	Age, race/ ethnicity, smoking, and BMI	- Subjects who participate in PA for leisure were less likely, as compared with sedentary subjects, to report 14 or more unhealthy days in one of the four measures of HRQoL. The odds ratio ranged from 0.40 (physical or physical limitations) to 0.57 (mental health). - The odds of unhealthy days was even lower among those who followed PA recommendations, ranging from 0.24 (physical limitations) to 0.48 (mental health)
Cassidy[40]	18 points (82/high)	Australia/not reported	278 women; over 70 years of age	Questionnaire developed for the study: exposure to PA defined as "inactive" ($<$ 3 hours/week) or "physically active" ($\geqslant$ 3 hours/week).	HRQoL and domains of SF-36	None	- General HRQoL scores in the "physically active" group were significantly higher (78.7 ± 13.7 vs. 66.5 ± 19.6, $p < 0.001$) than in the "physically inactive" group. - Physically active women were half as likely to be depressed and anxious when compared to their physically inactive counterparts (OR = 0.5, 95%CI 0.3-0.8 for both).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Brown[41]	21 points (95/high)	U.S./2001	175,850 subjects, among whom 30,146 were over 65 years of age	Questionnaire developed for the BRFSS to assess PA	Questionnaire developed for the BRFSS to assess QoL	Age, race/ ethnicity, gender, education, smoking, and BMI	- After adjustment for confounding factors, non-participation in moderate and vigorous PA was independently associated with increased odds of experiencing 14 or more unhealthy days (physical or mental), as compared with subjects who participated in moderate and vigorous PA (OR = 2.42 and 2.48 for moderate and vigorous PA, respectively) - Seniors who performed moderate PA with low weekly frequency (1-2 days/week) or daily (7 days/week) had a 30% greater chance of experiencing 14 or more unhealthy days, compared with those who underwent a schedule of moderate PA for 5-6 days/week. - Participation in moderate PA for 1-19 minutes/day (compared to a reference of 30-59 minutes/day) was also associated with a greater chance of having many unhealthy days (OR = 1.44)
Guallar-Castillón[42]	21 points (95/high)	Spain/2000-2001	3,066 subjects (54% women); over 60 years of age	Questionnaire developed for the study to assess the frequency of PA participation	HRQoL and domains of SF-36	Socio-demographic variables, lifestyle, use of health services, number of diseases, and social network	- In both sexes, sedentary subjects had worse scores on all scales of the SF-36 (linear trend $p < 0.05$). The main differences were observed in the Functional Capacity domain. - After adjusting for potential confounders, the practice of light PA was associated with higher scores on scales of the SF-36, except for physical and emotional domains, for both men and women (all $p < 0.05$). - In virtually all domains, the change in scores was greater than 3 points ($\beta > 3$). Compared to the sedentary group, participation in moderate/vigorous PA was associated with increased scores in functional capacity ($\beta > 15$), general health ($\beta > 9$), and vitality ($\beta > 9$), in both sexes, and with increased scores in physical aspects ($\beta = 13.2$), social functioning ($\beta = 6.13$), and mental health ($\beta = 9.9$) among men (all $p < 0.01$).
Lai[43]	19 points (86/high)	China/not reported	465 subjects (50.5% women); mean age 73.3±6.1 years	ECOG scale	General QoL and domains of WHOQoL-Bref	None	- Physical performance was independently associated with the physical domain of QoL ($r^2 = 0.40$); the lower the physical performance, the lower the physical domain scores. - Physical performance (ECOG scale) had the strongest impact on physical domain ($r^2 = 0.40$, $\beta =$ -0.62) and accounted for significant percentages of the variance on the psychological ($r^2 = 0.09$, $\beta = 0.30$), social relationships ($r^2 = 0.06$, $\beta = -0.26$), and environment ($r^2 = 0.13$, $\beta = -0.35$) domains (all p-values < 0.01).
Acree[44]	20 points (91/high)	U.S./not reported	112 subjects (56.6% women); between 60 and 89 years of age	Johnson Space Center Physical Activity Scale	HRQoL and domains of SF-36	Gender and hypertension	- All eight domains of HRQoL were higher ($p < 0.05$) in the active group. After adjustment for confounders, the domains of HRQoL that were significantly higher in the active group were physical function ($t = 4.90$, $p = 0.029$), physical ($t = 5.40$, $p = 0.022$), vitality ($t = 11.44$, $p =$ 0.001), social function ($t = 4.32$, $p = 0.04$), and pain ($t =$ 11.44, $p = 0.001$).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Brown[45]	21 points (95/high)	U.S./2003	215,670 subjects, among whom 42,368 were over 65 years of age	Questionnaire developed for the BRFSS	Questionnaire developed for the BRFSS to assess HRQoL	Age, gender, race/ ethnicity, education, smoking, BMI, and health problems	- After adjustment for confounders, hypertensive elderly who followed the PA recommendation were less likely to experience 14 or more unhealthy days in the five measures of HRQoL (physical and mental health, combination of both, physical problems, poor or reasonable health status), compared with inactive subjects (did not perform moderate/vigorous PA). The odds ratios were 0.33 (physical limitations), 0.37 (physical health), 0.38 (physical and mental health), 0.39 (poor or reasonable health status), and 0.45 (mental health). - Low levels of PA were associated with higher prevalence of 14 or more unhealthy days in the five measures of HRQoL in hypertensive elderly patients.
McAuley[46]	20 points (91/high)	U.S./2002-2003	249 women; between 59-84 years of age (mean age 68.1 years)	PASE CHAMPS	Physical health assessed with LL-FDI Mental health assessed with SF-12 General QoL assessed with SWLS	None	- PA was significantly related to physical health (β = 0.56), mental health (β = 0.19), and life satisfaction (β = 0.29). - In model 2, being physically active was associated with better mental (β = 0.29) and physical health (β = 0.56), which were positively associated with global QoL (β = 0.45 and 0.39, respectively). These results support the theory that physical and mental health are indicators of the relationship between PA and QoL. - In model 3, PA was directly associated with self-efficacy (β = 0.66). Self-efficacy was related to physical and mental health (β = 0.55 and 0.32, respectively), while the latter two were positively associated with global QoL (β = 0.42 and 0.41, respectively). Taken together, these findings support the idea that self-efficacy, physical health, and mental health have mediating roles in the relationship between PA and QV.
Lobo[47]	19 points (86/high)	Portugal/not reported	185 institutionalized subjects (51.3% women); over 60 years of age	Accelerometer (minutes/day of moderate or vigorous PA)	HRQoL and domains of SF-36	None	- Active elderly women had significantly higher values of HRQoL in physical functioning (r = 0.351) and pain (r = -0.719). Additionally, among men, there was also a significant correlation of moderate PA with physical functioning (r = 0.531), physical health (r = 0.311), vitality (r = 0.219), pain (r = -0.821), and mental health (r = 0.252). - Among men, total PA (counts/day) significantly correlated with five variables of HRQoL, negatively with pain (r = -0.632) and positively with the other domains: physical functioning (r = 0.395), physical health (r = 0.268), vitality (r = 0.229), and mental health (r = 0.230). Among women, total PA was negatively associated with pain (r = -0.502).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Mazo[2]	18 points (82/high)	Brazil/not reported	198 women; over 65 years of age	IPAQ – long version	General QoL and WHOQoL-Bref domains	None	- Active elderly women had higher average scores compared to their less active peers in the physical (3.7±0.6 vs. 3.4±0.5, $p < 0.01$) and psychological domain (3.8±0.5 vs. 3.7±0.5, p = 0.02). No difference was found between the other domains, social relations (4.1±0.4 vs. 4.0±0.4, p = 0.372) and environment (3.4±0.4 vs. 3.4±0.5, p = 0.794). - The results indicate significant associations between the physical domain and the PA level (less and more active); hence, elderly that have a poorer outcome in the physical domain showed a higher risk of being less physically active (OR = 2.9, 95%CI 1.4-5.5, p = 0.010). The other domains did not show significant associations (psychological, OR = 1.6, 95%CI 0.7-3.5, p = 0.240; social relations, OR = 0.9, 95%CI 0.4-2.0, p = 0.812; environment, OR = 0.5, 95%CI 0.2-1.3, p = 0.178; general QoL domain, OR = 1.0, 95%CI 0.6-1.8, p = 0.981).
Rossato[48]	19 points (86/high)	Brazil/2006-2007	103 subjects (72.8% women); over 60 years of age (mean age 70.8 years)	Barthel Index, ranking individuals as independent or dependent	General QoL and WHOQoL-Bref domains	None	- There were significant differences between the independent and dependent groups in all domains of QoL and general QoL, with higher values among the independent ($p < 0.05$): physical (15.10±2.72 vs. 13.50±2.88, p = 0.05), psychological (15.60±1.83 vs. 14.22±2.85, p = 0.04), social relations (16.14±1.82 vs. 14.35±2.62, p = 0.000), environment (15.16±1.49 vs. 14.01±2.28, p = 0.03), and general QoL domain (15.60±2.21 vs. 13.84±3.30, p = 0.02).
Johnson[49]	21 points (95/high)	U.S./2006-2007	843 subjects with colorectal cancer (53% women); between 65-84 years	CHAMPS	HRQoL and SF-36 domains	Gender, age, educational attainment, BMI, pain, recent hospitalization, falls, and comorbidities	- After adjustment for confounders, the more active individuals (4th quartile) had higher scores in the physical domain, compared with less-active individuals (1st quartile) (66.0 vs. 42.7). - PA participation was associated with better function in the physical domain (p = 0.004).
Alexandre[50]	19 points (86/high)	Brazil/2005-2006	120 subjects (85.8% women); between 60-89 years of age (mean age 69, standard deviation = 5)	Six-minute test Mobility test for Timed Up and Go Test HAP	General QoL assessed with WHOQoL-Bref	None	- Higher weekly frequency of PA was associated with higher average values in the physical domain (61.71 vs. 65.32, p = 0.028), but not in other domains (psychological = 61.33 vs. 64.65, p = 0.292; social = 72.66 vs. 73.65, p = 0.640; environment domain = 65.06 vs. 70.86, p = 0.077). - There were no differences in the means of different groups for overall physical fitness and balance. - The results of the independent correlations for each domain regarding QoL of active elderly showed a positive association in all domains (physical, β = -0.77, VIF = 3.30; psychological, β = -1.18, VIF = 3.15; social, β = -1.51, VIF = 3.38; environment domain = β = -1.24, VIF = 3.28, all $p \leq 0.001$).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Ekwall[51]	19 points (86/high)	Sweden/not reported	1,924 subjects (63.4% women) with history of dizziness; over 75 years of age	Questionnaire developed for the study to assess the frequency of light and vigorous PA participation	HRQoL and SF-12 domains	Gender, age, health problems, and frequency of falls	- After adjustment for confounding factors, the practice of light PA was associated with 40% lower odds of having low HRQoL (physical or mental). The practice of vigorous PA was associated with 50-70% lower odds of having lower HRQoL (physical and mental). - Participation in extensive PA presented a significant association ($p < 0.001$) with physical (OR = 0.206, 95%CI 0.124-0.341) and mental domains (OR = 0.478, 95%CI 0.311-0.736), the same was found to light PA ($p = 0.003$) in physical (OR = 0.615, 95%CI 0448-0.843) and mental (OR = 0.583, 95%CI 0.410-0.828) domains.
Paskulin[52]	21 points (95/high)	Brazil/2000	288 subjects (67.4% women); over 60 years of age (mean age 71.2 years)	The Older American Resources and Services	General QoL and domains of WHOQoL-Bref	Age, gender, educational attainment, health problems, and behavioral factors	- After adjusting for confounders, being involved in PA was associated with a seven-point increase in physical domain and global QoL scores ($\beta = 0.18$ and $\beta = 0.19$, respectively, $p < 0.01$). - High levels of dependency for daily PA were associated with a reduction of about 11-13 points in physical domain scores ($\beta = 0.14$, $p < 0.01$), and severe dependence was associated with a 15-point reduction in the psychological domain scores ($\beta = 0.21$, $p < 0.01$). In addition, the mild level of dependence was significantly associated with a five-point reduction in psychological domain scores ($\beta = 0.13$, $p = 0.03$).
Renne-mark[53]	22 points (100/high)	Sweden/ 2001-2003	1,005 subjects (56.5% women); mean age 75.3 years	Moderate and vigorous PA assessed by two questions: "How often have you been involved in easy exercises for the past 12 months?" and "How often have you exercised with higher intensity for the past 12 months?"	Wellness and personal relationship were measured by two components extracted from the LGC questionnaire. Emotional support and social anchorage were measured in terms of two factors of social networks taken from a previously developed questionnaire	Age, gender, functional ability, and comor-bidities	- After adjustment for confounding factors, being involved in light PA was significantly associated with an increase of 11-21% in the chance to experience well-being (OR = 1.15, 95%CI 1.01-1.30), social engagement (OR = 1.21, 95%CI 1.08-1.36) and social encouragement (OR = 1.11, 95%CI 1.00-1.24). Additionally, participation in vigorous PA was significantly associated with a 15% increase in the chance to experience social engagement (OR = 1.15, 95%CI 1.03-1.28) and emotional support (OR = 1.15, 95%CI 1.03-1.29).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Toscano[54]	18 points (82/high)	Brazil/not reported	238 women; mean age 69.2 years	IPAQ – long version	General QoL and SF-36 domains	None	- Active elderly women demonstrated better results in the eight QoL domains ($p < 0.05$) when compared to their less active peers (functional capacity = 70.1±20.1 vs. 46.9±23.4; physical aspects = 75.3±35.4 vs. 52.9±37.9; pain = 64.1±23.1 vs. 52.2±26; general health = 69.1±18.4 vs. 48.8±20.1; vitality = 72.9±16 vs. 56.9±18.9; social aspects = 85.1±22.6 vs. 72.5±23; emotional aspects = 81.6±33.4 vs. 63.5±42.9; mental health = 56.6±7.3 vs. 51.8±7.7). - Greater differences were observed for functional capacity and physical aspects.
White[55]	19 points (86/high)	U.S./not reported	321 subjects (80.1% women); over 50 years of age (mean age 63.8 years)	GLTEQ; PASE	Physical health assessed with LL-FDI Exercise self-efficacy assessed with Exercise Self-Efficacy Scale Physical self-esteem assessed with Physical Self-Perception Profile General QoL assessed with SWLS	Gender, race, age, education, and income	- PA was significantly associated with self-efficacy for exercise ($\beta = 0.60$), which, in turn, was associated with low limitations/deficiencies ($\beta = 0.28$), and greater physical self-esteem ($\beta = 0.44$), low limitations/ deficiencies ($\beta = 0.20$), and high physical self-esteem ($\beta = 0.40$) were significantly associated with better general QoL. - The correlation between PA and the model constructs showed significant values (exercise self-efficacy, $r = 0.28$, $p < 0.01$; disability limitations, $r = 0.15$, $p < 0.1$; physical self-esteem, $r = 0.17$, $p < 0.01$), except with the satisfaction with life ($r = 0.05$, $p > 0.05$). - These observations support the socio-cognitive theory concept that self-efficacy for exercise and physical and mental health play an intermediary role in the relationship between PA and QoL. After adjusting for confounding factors, the coefficients did not change drastically; however, the beta values of the relationship between PA and self-efficacy changed from 0.60 to 0.73
Carvalho[56]	20 points (91/high)	Brazil/2009-2010	271 women; over 60 years of age	IPAQ – long version, adapted for elderly	General QoL and domains of WHOQoL-Old	None	- Those who participated in moderate/vigorous PA had higher QoL scores in almost all domains when compared with those who did not participate: sensory ability (77.1±16.9 vs. 67.7±19.8), autonomy (63.3±15.7 vs. 56.6±16.0), past, present, and future activities (71.9±15.2 vs. 64.3±17.4), social participation (71.8±15.1 vs. 62.3±17.4), and intimacy (70.3±17.0 vs. 65.3±18.8, all $p < 0.01$), with the exception of the death and dying domain (65.2±23.8 vs. 64.4±22.5, $p = 0.59$). - Practicing moderate/vigorous PA increased general QoL scores by 4.8±1.5 points.

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Salguero[57]	19 points (86/high)	Spain/not reported	436 subjects (53.6% women), institutionalized and from the community; between 60 and 98 years of age	YPAS	HRQoL and domains of SF-36 GDS	PA level, gender, and institutionalization	- With the exception of the emotional aspects domain, all domains of HRQoL were significantly correlated with indicators of PA. However, we observed low correlation values (all r-values < 0.31). The HRQoL domains that had significant correlations with total PA were: physical function (r = 0.192), role physical (r = 0.139), physical appearance (r = 0.168), general health (r = 0.109), and vitality (r = 0.118). - The index of vigorous PA was significantly correlated with physical function (r = 0.259), role physical (r = 0.152), bodily pain (r = 0.120), physical appearance (r = 0.210), general health (r = 0.178), vitality (r = 0.159), mental health (r = 0.154), and depressive symptoms (r = 0.173). - Among the institutionalized, almost all domains had significantly higher scores among the most active, with the exception of the social function domain, regardless of gender. Among the community residents, the physical function domain was higher among active individuals, in both sexes. - The general health domain among men (community dwellers, 69±17 vs. 59±18; institutionalized, 63±30 vs. 52±23) and the physical appearance (community dwellers, 49±7 vs. 45±8; institutionalized, 49±8 vs. 39±11) and pain (community dwellers, 75±20 vs. 68±22; institutionalized, 81±24 vs. 59±34) domains among women were higher in the active elderly, compared to their less active peers.
Figueira[58]	18 points (81/high)	Brazil/not reported	100 subjects (92% women); mean age 60 years	IPAQ - long version	SF-36 domains	None	- Active elderly showed higher scores in the pain(41.5±26.1 vs. 35±15.6), physical aspects (37.9±49.1 vs. 9.2±24.1) and functional capacity (44.7±32 vs. 21.5±20.9) domains when compared with the sedentary elderly.
Grimmett[59]	19 points (86/high)	England/2005	478 subjects (40.58% women); mean age 68 years	GLTEQ	EORTC-QLQ-C30	Demographic (age, sex, socioeconomic status) and medical covariates (comorbidities, recurrence, current treatment and time since diagnosis)	- Elderly that practice $\geqslant 5$ sessions per week presented higher scores than that < 5 sessions per week in the following domains: global QoL (77.34±2.25 vs. 70.36±1.30), physical function (90.11±1.48 vs. 81.77±0.98), role function (87.87±2.36 vs. 81.82±1.43), social function (83.39±2.98 vs. 76.56±1.46).
Guedes[60]	20 points (91/high)	Brazil/2009	1,204 subjects (53.57% women); aged $\geqslant$ 60 years	IPAQ – short version	WHOQoL-Old	Age, marital status, educational level, and family socioeconomic status.	- Active elderly showed significantly greater values when compared to the sedentary in the following domains: sensory abilities (76.39±17.89 vs. 71.08±18.72), autonomy (62.13±15.21 vs. 54.64±16.07), social participation (74.32±15.72 vs. 66.79±16.84), intimacy (53.84±16.34 vs. 48.17±17.95), and overall QoL (62.27±10.15 vs. 57.15±11.26).

Continued on next page

Supplementary Table Continued

Author	Quality assessment score* (%/classification)	Country/ date of collection	Method of sample selection; sample size; age	PA programs	Measurement of QoL	Adjustment variables	Main findings
Lima[61]	20 points (91/high)	Brazil/2001-2002	1,958 subjects (53.55% women); mean age of 69.6 years	Questionnaire developed for the study to assess the level of PA participation	SF-36 domains	Gender, age, educational level, income, work status, area of residence, and number of chronic conditions	- PA was positively associated with SF-36 subscales: role physical (91.5, 95%CI 83.4-99.7, β = 11.9), physical functioning (82.2, 95%CI 76.1-88.3, β = 11.3), bodily pain (79.7, 95%CI 74.1-85-3, β = 4.5), general health (75.6, 95%CI 71.2-80.1, β = 5.6), vitality (69.5, 95%CI 63.7-75.1, β = 4.1), role emotional (94.3, 95%CI 87.8-100, β = 9.9), social functioning (93.5, 95%CI 86.8-100, β = 8.6), and mental health (73.5, 95%CI 69.3-77.7, β = 2.8).
Vala-dares[62]	21 points (95/high)	Brazil/not reported	271 women; mean age 77.4 years	IPAQ – long version	WHOQoL-Old	Place of origin, age, number of medications used, social class, BMI, smoking, marital status, educational level, perception of health, history of disease, and paid employment	- The past, present, and future activities (β = 0.021) and social participation (β = 0.03) domains were associated with the IPAQ transportation section. - The leisure-time domain of the IPAQ had an inverse and significant relationship with the QoL score related to past, present, and future activities (β = -0.0269).
Porto[63]	19 points (86/high)	Brazil/2006	199 subjects (41.21% women); aged between 60 to 70 years	IPAQ – short version	WHOQoL-Bref	Gender, schooling, and income	- The physical domain presented the strongest association with physical inactivity (OR = 2.70, 95%CI 1.87-3.52), and the other domains also presented greater odds of being in the lowest QoL scores (psychological OR = 2.45, 95%CI 1.45-3.11; social relationships OR = 2.14, 95%CI 1.19-3.26). - Only the environmental domain was not associated with physical inactivity.
Baern-holdt[64]	21 points (95/high)	U.S./2005-2006	911 subjects (56.8% women); aged between 65 to 85 years	Andersen's behavioral model	Domains of SF-36 and WHOQoL-Bref	Age, gender, race, family size, marital status, education, and poverty index	- PA was linked to social functioning ($p > 0.05$), and emotional well-being was lower in people with less PA (estimate -0.045 and regression coefficient = 0.42).

95%CI = 95% confidence interval; ACSM = American College of Sports Medicine; BMI = body mass index; BRFSS = Behavioral Risk Factor Surveillance System; CG = control group; CHAMPS = Community Healthy Activities Model Program for Seniors physical activity measure; ECOG scale = Eastern Cooperative Oncology Group scale; EORTC-QLQ-C30 = Quality of life was assessed with the European Organization for Research and Treatment of Cancer Quality of Life questionnaire, version C30; FES-I = Falls Efficacy Scale; GDS = Geriatric Depression Scale; GLTEQ = Godin Leisure Time Exercise Questionnaire; HAP = Human Activity Profile; HRQoL = health-related quality of life; IADL = Instrumental Activities of Daily Living Scale; IG = intervention group; IPAQ = International Physical Activity Questionnaire; LGC questionnaire = Lund Gerontology Centre Life Quality Questionnaire; LL-FDI = Late-Life Function and Disability Instrument; MET = metabolic equivalent; MOS = Medical Outcomes Study; MUNSH = Memorial University of Newfoundland Scale of Happiness; NPI = Neuropsychiatric Inventory; OR = odds ratio; PGC morale scale = Philadelphia Geriatric Center Morale Scale; PA = physical activity; PASE = Physical Activity Scale for the Elderly; PSW = Physical Self-Worth scale; QoL = quality of life; SF-12 = SF-36 Physical and Mental Component Summary; SF-36 = Short Form-36; SWLS = Satisfaction with Life Scale; TMIG-IC = Tokyo Metropolitan Institute of Gerontology Index of Competence; VIF = variance inflation factor; WHOQoL-100 = World Health Organization Quality of Life 100; WHOQoL-Bref = World Health Organization Quality of Life Assessment – Abbreviated Version; WHOQoL-Old = World Health Organization Quality of Life Assessment – Module for Older Adults; YPAS = Yale Physical Activity Questionnaire.
* The Downs & Black[21] checklist was used for intervention studies (with scores ranging from 0 to 27 points), and an adaptation of the Strengthening the Reporting of Observational Studies in Epidemiology-STROBE checklist, for cohort and cross-sectional studies (with scores ranging from 0 to 22 points).