Supplementary Table 2. Summary of Studies Included in the Meta-Analysis

Author	Country	Number of Odds Ratios (Crude/ Adjusted)	Study Design	Sample Size (% female)	Participants	Age (years)	Follow-up (years)	PA Measure	PA Parameters	Depression Measure	Depression Method
Abjibade et al, (2018)	France	1/0	Baseline exposure	25837 (76.1)	Participants from NutriNet-Santé study	47.3	5.40	IPAQ	Volume	CES-D ≥ 17 (men) and ≥ 23 (women)	Screening level
Amara et al, (2019)	USA	0/1	Change in outcome	380 (34.5)	Participants from the PPMI study	63.55	2.00	PASE	Volume	GDS-15 (continuous)	Symptoms
Anagnostopoulos et al, (2015)	Switzerland	2/0	Baseline exposure	6756 (27.3)	Participants from Swiss HIV Cohort Study	NR	3.00	Single question on frequency of exercise	Frequency	Psychiatrist diagnosis	Clinical diagnosis
Andrade-Gomez et al, (2018)	Spain	0/4	Change in outcome	1461 (51.1)	Participants from the Seniors-ENRICA cohort	70.77	3.00	Validated questionnaire	MET.hours	GDS-10 (continuous)	Symptoms
Augestad et al, (2008)	Norway	0/6	Baseline exposure	6661 (50.3)	Participants from HUNT Cohort Study	21-40	11.00	Questions on frequency, duration, and intensity of PA	Volume	HADS-D ≥ 8	Screening level
Backmand et al, (2009)	Finland	4/0	Baseline exposure	853 (0)	Former elite male athletes and military	52.6	10.00	Three questions on PA intensity, duration, and frequency	MET.hours	Highest 10% of scores in short stress symptom survey	Screening level
Ball et al, (2009)	Australia	4/4	Baseline exposure Change in outcome	6677 (100)	Participants from ALSWH	22-27	3.00	Questions on PA frequency and duration	MET.minutes	CES-D ≥ 10	Screening level
Baumeister et al, (2017)	Germany	0/2	Baseline exposure	1,043 (49.7)	Participants from SHIP Study	20-79	4.50	Baecke questionnaire	Volume	BDI-II ≥ 12	Screening level
Beard et al, (2007)	Australia	3/0	Baseline exposure	1407 (56.0)	Participants from Northern Rivers Mental Health Study	18+	2.00	Questions on PA type, frequency, and duration	Time	MiniCIDI (DSM-IV based interview)	Clinical diagnosis
Bernaards et al, (2006)	Netherlands	0/6	Change in exposure	1747 (57.2)	Participants from SMASH	36.0	3.00	One question on PA frequency	Frequency	CES-D-11 ≥ 6	Screening level
Bots et al, (2007)	Finland, Italy, Netherlands	1/1	Baseline exposure Change in outcome	526 (0)	Participants from FINE Study	75.0	5.00	Morris questionnaire	Frequency & time	ZSDS ≥ 48	Screening level
Boyes et al, (2013)	Australia	0/1	Change in outcome	1154 (41.6)	Cancer patients	18-80	0.50	Three questions on PA frequency and	Volume	HADS-D ≥ 8	Screening level

							duration				
Branstrom et al, (2015)	Sweden	0/1	Change in exposure and outcome	726 (100)	Breast cancer patients	51.3	2.00	Single question on PA frequency	Frequency	HADS-D $\geq$ 8	Screening level
Brown et al, (1996)	US	1/0	Baseline exposure	1322 (55.0)	Participants from NHANES	20-79	8.00	Two questions on recreational and non-recreational PA	Time	CES-D $\geq$ 16	Screening level
Brown et al, (2005)	Australia	0/2	Change in exposure and outcome	8855 (100)	Participants from ALSWH	47.5	4.00	Questions on PA frequency and duration	Volume	CES-D-10 $\geq$ 10	Screening level
Brunet et al, (2018)	Canada	2/0	Change in exposure Change in exposure and outcome	201 (100)	Women post-breast cancer treatment in Montreal	55.0	1.50	Actigraph GT3X accelerometer	Volume	CES-D-10 (continuous)	Symptoms
Byers et al, (2012)	US	0/1	Baseline exposure	7240 (100)	Participants from Study of Osteoporotic Fractures	72.8	12.20	Modified Paffenbarger Activity Questionnaire	Volume	GDS (continuous)	Symptoms
Cabello et al, (2017)	Ghana, India, Mexico, and Russia	1/0	Baseline exposure	5970 (66.5)	Participants from World Health Organization's SAGE	50.2	5.00	IPAQ short form	MET.hours	SR of clinical diagnosis, treatment, or ICD-10 depression	SR diagnosis
Camacho et al, (1991)	US	0/2	Change in exposure Change in exposure and outcome	4848 (55.6)	Participants from Alameda County Study	50.0	9.00	Question on PA type and frequency	Frequency	Alameda County Study scale $\geq$ 5	Screening level
Capella McDonnall, (2011)	US	1/0	Change in exposure and outcome	2688 (54.1)	Elderly people with dual sensory loss	69.2	13.00	One question on PA frequency	Frequency	CES-D-8 (continuous)	Symptoms
Carroll et al, (2010)	USA	0/2	Change in exposure and outcome	5389 (57.2)	Participants from Health and Retirement Study	55.5	6.00	Single question on PA frequency	Frequency	Male: CES-D $\geq$ 4 Female: CES-D $\geq$ 3	Screening level
Chang et al, (2015)	Iceland	0/1	Baseline exposure	4140 (57.0)	Participants from Reykjavik Study	52.0	25.00	Two questions on PA frequency and duration	Time	GDS-15 $\geq$ 6	Screening level

Chang et al, (2017)	Taiwan	0/4	Change in outcome	2673 (45.5)	Participants from TLSA	74.2	8.00	Questions on frequency, duration, and intensity of PA	Frequency	CESD ≥ 10	Screening level
Chao, (2014)	Taiwan	0/2	Change in exposure and outcome	2660 (43.9)	Participants from SHLSE	80.8	3.00	One question on frequency of PA	Frequency	CES-D (continuous)	Symptoms
Chen & Millar, (1999)	Canada	1/1	Baseline exposure	7593 (53.8)	Participants from National Population Health Survey	38.0	2.00	Questions on PA type, frequency, and duration	Kilocalories/ kilogram of body weight/ day	DSM III-R CIDI	Diagnosis
Chi et al, (2016)	Taiwan	0/2	Baseline exposure	2630 (46.8)	Participants from TLSA	53+	4.00	Questions on frequency, duration, and intensity of PA	Volume	CES-D ≥ 10	Screening level
Choi et al, (2015)	Korea	1/1	Baseline exposure	5327 (52.6)	Participants from KLoSA	45+	2.00	Questions on exercise frequency and duration	Time	CES-D-10 ≥ 4	Screening level
Choi et al, (2020)	UK	0/1	Baseline exposure	7968 (57.0)	Participants from UK Biobank	60.0	2.00	Series of questions on time engaged in eight types of PA	MET.hours	ICD-9 or 10 diagnosis	Diagnosis
Chuang et al, (2012)	USA	0/1	Change in outcome	1404 (100)	Participants from CePAWHS	18-45	2.00	Single question: MVPA ≥ 30 min most days of week	Volume	Modified CES-D ≥ 4	Screening level
Cooper-Patrick et al, (1997)	USA	2/2	Baseline exposure	752 (8.0)	Participants from Precursors Study	50.0	15.00	Harvard Alumni PA questionnaire	Frequency	SR of clinical depression	SR diagnosis
Da Silva et al., (2012)	UK	1/1	Baseline exposure	9309 (31.5)	Participants from Whitehall II Study	35-55	3.00	Two questions on PA duration and intensity	Volume	30-item GHQ ≥ 4	Screening level
Dugan et al, (2015)	USA	0/2	Change in outcome	2682 (100)	Participants from SWAN Study	45.9	10.00	Kaiser PA Survey	Volume	CES-D ≥ 16	Screening level
Edwards et al, (2016)	USA	0/1	Baseline exposure	489 (82.2)	People with multiple sclerosis, spinal cord injury, muscular dystrophy, or postpolio syndrome	54.3	3.50	GLTEQ	Volume	PHQ-9 ≥ 10	Screening level
Ensari et al, (2014)	USA	0/1	Change in outcome	269 (82.9)	People with multiple sclerosis	45.9	2.50	GLTEQ	Volume	HADS-D ≥ 8	Screening level
Ernstsen et al, (2016)	Norway	0/2	Change in exposure	189 (24.3)	Participants from HUNT Cohort Study	61.6	22.00	Questions on exercise frequency, intensity, and duration	Volume	HADS-D ≥ 8	Screening level

España-Romero et al, (2013)	USA	0/2	Baseline exposure	5110 (20.4)	Participants from Aerobic Center Longitudinal Study	47.5	6.10	Questions on PA frequency, intensity, and duration	MET.minutes	CES-D-10 ≥ 8	Screening level
Farmer et al, (1988)	USA	0/2	Baseline exposure	1163 (51.5)	Participants from NHANES	50.0	9.00	Questions on recreational and nonrecreational PA	Frequency	CES-D ≥ 16	Screening level
Ford et al, (1998)	USA	1/0	Change in exposure and outcome	1190 (0)	Participants from Precursors Study	66.0	37.00	Harvard Alumni PA questionnaire	Frequency	SR of clinical depression	SR diagnosis
Foreyt et al, (1995)	USA	2/0	Change in exposure and outcome	381 (48.9)	Participants from Reno diet-heart study	44.1	4.00	Composite measure of SR recreational PA and PA importance	Frequency	CES-D (continuous)	Symptoms
Fukukawa, (2004)	Japan	0/1	Change in outcome	314 (41.7)	Participants from NILSLSA	65-79	2.00	Pedometer	Distance	CES-D (continuous)	Symptoms
Gallegos-Carrillo et al, (2013)	Mexico	2/2	Change in exposure	1047 (77.5)	Participants from Health Worker Cohort Study	18+	6.00	Questionnaire on PA duration and intensities	MET.hours	CES-D ≥ 16	Screening level
Garcia-Pena et al, (2013)	Mexico	0/1	Change in outcome	7882 (61.2)	Participants from Integrated Study of the Elderly	71.0	2.00	One question assessing exercise frequency	Frequency	GDS-30 ≥ 11	Screening level
Giltay et al, (2006)	Netherlands	1/0	Baseline exposure	464 (0)	Participants from Zutphen Elderly Study	70.8	15.00	Zutphen PA Questionnaire	Volume	ZSDS ≥ 50	Screening level
Gudmundsson et al, (2015)	Sweden	1/1	Change in outcome	676 (100)	Participants from PPSW	53.4	32.00	Saltin–Grimby PA Level Scale	Volume	MADRS (continuous)	Symptoms
Hallgren et al, (2019)	Sweden	2/0	Baseline exposure	25520 (64.8)	Participants from the Swedish National March Cohort study	18+	13.00	Nine questions on time engaged in physical activities of different intensities	Volume	MDD diagnosis made by a specialist clinician	Diagnosis
Hamer et al, (2009)	England	0/2	Change in outcome	4323 (52.5)	Participants from ELSA	63.4	4.00	Questions on frequency of moderate and vigorous PA	Frequency	CES-D-8 ≥ 4	Screening level
Han et al, (2019)	South Korea	8/0	Baseline exposure	2502998 (60.7)	Adults in South Korea	20+	1.00	Reported days per week engaged in specific doses of PA	Frequency	Korean Classification of Diseases diagnosis	Diagnosis
Harvey et al, (2017)	Norway	0/1	Baseline exposure	33908 (49.5)	Participants from HUNT Cohort Study	45.2	11.00	Questions on frequency, duration, and	Time	HADS ≥ 8	Screening level

							intensity of exercise				
Henchoz et al, (2014)	Switzerland	4/0	Baseline exposure Change in exposure and outcome	4846 (0)	Participants from C-SURF	20.0	1.42	One question on exercise frequency	Frequency	MDI (continuous)	Symptoms
Herbolsheimer et al, (2018)	Germany	0/1	Change in outcome	334 (39.2)	Participants from the from the ActiFE Ulm study	72.6	3.00	activPAL accelerometer	Time	HADS-D (continuous)	Symptoms
Hiles et al, (2015)	Australia	0/2	Baseline exposure	1410 (50.4)	Participants from Hunter Community Study	65.6	4.50	Pedometer	Distance	CES-D ≥ 16	Screening level
Hiles et al, (2017)	Australia	0/1	Change in outcome	2932 (66.4)	Participants from NESDA	41.9	2.00	IPAQ short form	MET.minutes	IDS (continuous)	Symptoms
Hurwitz, (2003)	USA	0/1	Baseline exposure	2902 (53.0)	Participants from CWHS	18+	1.00	Question on PA frequency	Frequency	GDS-short form ≥ 7	Screening level
Ivanova et al, (2016)	Canada	2/2	Change in outcome	1691 (49.1)	Participants from Evaluation of Diabetes Treatment study	63.9	1.00	One question on exercise frequency	Frequency	PHQ-9 (continuous)	Symptoms
Jantunen et al, (2019)	Finland	0/2	Change in exposure and outcome	890 (57.2)	Participants from the Helsinki Birth Cohort Study	61.3	10.00	Validated questionnaire	MET.hours	BDI (continuous)	Symptoms
Joshi et al, (2016)	USA	0/3	Change in outcome	2023 (60.8)	Participants from NYCNAMES	70.2	1.00	PASE	Volume	PHQ-9 ≥ 10	Screening level
Kanamori et al, (2018)	Japan	2/2	Baseline exposure	1422 (45.9)	Participants from JAGES	72.5	2.00	Two questions on exercise frequency	Frequency	GDS-15 ≥ 5	Screening level
Kaseva et al, (2016)	Finland	0/1	Change in outcome	1724 (50.9)	Participants from Cardiovascular Risk in Young Finns Study	41.4	5.00	Questions on PA frequency, intensity, and duration	Frequency	BDI-II (continuous)	Symptoms
Khalaila & Litwin (2014)	Israel	0/2	Change in exposure and outcome	1524 (57.0)	Participants from Israeli component of SHARE	50+	4.00	Two questions on moderate and vigorous PA	Frequency	EURO-D ≥ 4	Screening level
Khalaila, (2016)	Israel	1/1	Baseline exposure	1038 (58.2)	Participants from Israeli component of SHARE	50+	10.00	Two questions on moderate and vigorous PA	Frequency	EURO-D ≥ 4	Screening level
Knox et al, (2006)	USA	2/2	Change in exposure	5115 (54.4)	Participants from CARDIA	35.5	15.00	Interview administered questions on PA	METs	CES-D ≥ 16	Screening level

Study	Country	Quality	Design	N (%)	Population	Age	Follow-up	PA measure	Metric	Depression measure	Assessment
								type, frequency, and duration			
Kritz-Silverstein et al, (2001)	USA	0/4	Baseline exposure	944 (57.2)	Participants from Rancho Bernardo Heart and Chronic Disease Study.	70.4	8.00	Two questions on strenuous exercise and labour frequency	Frequency	Modified BDI ≥ 13	Screening level
Ku at al, (2009)	Taiwan	2/2	Baseline exposure Change in exposure	2831 (45.0)	Participants from Longitudinal Survey of Health & Living status of the Elderly, Taiwan	50+	7.00	One question on PA frequency	Frequency	CES-D-10 ≥ 10	Screening level
Ku at al, (2012)	Taiwan	2/2	Change in outcome Change in exposure and outcome	1160 (49.4)	Participants from Longitudinal Survey of Health & Living status of the Elderly, Taiwan	67+	11.00	One question on PA frequency	Frequency	CES-D-10 (continuous)	Symptoms
Lampinen et al, (2000)	Finland	0/1	Change in exposure and outcome	663 (64.0)	Participants from Evergreen project	75.0	8.00	Question on physical exercise type	Frequency	Modified BDI-13 ≥ 5	Screening level
Lee et al, (2017)	Australia	1/0	Change in outcome	218 (58.0)	Post-hospitalized older adults	65+	0.50	Phone-FITT household and recreational subscales	Time	GDS-short form ≥ 6	Symptoms
Lindwall et al, (2011)	Europe	1/0	Baseline exposure	17593 (54.6)	Participants from SHARE	64.1	2.00	Two questions on PA frequency and intensity	Frequency	Affective suffering factor of EURO-D (continuous)	Symptoms
Lindwall et al, (2014)	Sweden	1/0	Change in exposure and outcome	3717 (86.0)	Swedish health care workers	46.9	6.00	Saltin–Grimby PA Level Scale	Volume	HADS-D (continuous)	Symptoms
Lucas et al, (2011)	USA	4/0	Baseline exposure	49821 (100)	Participants from Nurses' Health Study	63.0	2.00	SR average time spent in various activities per week	Volume	SR physician diagnosis or regular use of antidepressants	SR diagnosis
McDowell et al, (2018)	Ireland	3/3	Baseline exposure	4146 (44.6)	Participants from TILDA	50+	2.00	IPAQ short form	MET.hours	CES-D ≥ 16	Screening level
Messier et al, (2013)	Canada	1/2	Baseline exposure Change in exposure	1183 (53.6)	Participants from Montreal Diabetes Health and Well-Being Study	18-80	1.00	One question on PA frequency	Frequency	PHQ-9 ≥ 5	Screening level
Mihrshahi et al, (2014)	Australia	3/3	Baseline	5117	Participants from	55.5	3.00	Questions on PA	MET,minutes	CES-D-10 ≥ 10	Screening

			exposure	(100)	ALSWH			frequency and duration			level
Mobily et al, (1996)	USA	0/2	Baseline exposure	2084 (63.2)	Participants from Rural Health Study	73.2	3.50	One question on walking frequency	Frequency	CES-D ≥ 16	Screening level
Monin et al, (2015)	USA	0/2	Change in exposure and outcome	2520 (50.0)	Participants from Cardiovascular Health Study	72.4	7.00	Questions on PA type, frequency, and duration	Kilocalories/ week	Modified CES-D (continuous)	Symptoms
Morgan & Bath, (1998)	UK	0/1	Change in outcome	690 (57.2)	Participants from NLSAA	75.6	4.00	Questions on PA frequency and duration	Time	SAD ≥ 6 & SAD-D ≥ 4	Screening level
Motl et al, (2012)	USA	1/1	Change in outcome	218 (93.7)	People with multiple sclerosis	43.5	1.50	IPAQ short form	METs	HADS-D (continuous)	Symptoms
Oh et al, (2020)	Korea	2/2	Baseline exposure	3873 (57.3)	Participants from KLOSCAD	70.4	4.00	?	MET.minutes	MINI-K diagnosis	Diagnosis
Ohrnberger et al, (2017)	UK	0/1	Change in outcome	10693 (55.0)	Participants from ELSA	66.3	2.00	Questions on frequency of moderate and vigorous PA	Frequency	CES-D (continuous)	Symptoms
Paffenbarger et al, (1994)	USA	0/2	Baseline exposure	10201 (0)	Harvard Alumni	54.5	25.00	Questions on PA types, frequencies, and durations	Volume	SR of physician diagnosis of depression	SR diagnosis
Park et al., (2015)	Korea	0/2	Baseline exposure	701 (59.9)	Participants from YEDD Study	72.2	5.00	Modified IPAQ	MET.hours	GDS ≥ 8	Screening level
Perrino et al, (2010)	USA	0/2	Baseline exposure Change in exposure and outcome	273 (59.0)	Participants from Hispanic Elders' Behavioral Health Study	70-100	1.00	SR walking routes during the preceding week	Distance	CES-D (continuous)	Symptoms
Pinto Pereira et al, (2014)	UK	0/3	Change in exposure and outcome	11135 (56.0)	Participants from 1958 British Birth Cohort	23.0	27.00	One question on PA frequency	Frequency	Malaise Inventory (top 10%)	Screening level
Poole & Jackowska (2018)	UK	0/1	Change in exposure	5172 (54.6)	Participants from ELSA	50+	4.00	Questions on frequency of moderate and vigorous PA	Frequency	CES-D (continuous)	Symptoms
Quezada et al, (2017)	Mexico	0/1	Change in outcome	456 (100)	Adult women of Health Workers Cohort	45.8	5.00	SR time per week spent in various activities within past year	MET.hours	CES-D (continuous)	Symptoms
Raudsepp & Riso (2017)	Estonia	0/2	Change in exposure	195 (56.4)	Community-dwelling older	72.1	1.00	Yamax-Digiwalker pedometer	Distance	GDS-15 (continuous)	Symptoms

			and outcome		adults						
Ribeiro et al, (2017)	USA	1/1	Change in outcome	550 (65.0)	Participants from African American Health Study	64.1	2.00	Yale Physical Activity Survey	Volume	CES-D-11 ≥ 9	Screening level
Rius-Ottenheim et al, (2013)	Netherlands	1/1	Change in exposure and outcome	600 (19.0)	Participants from Alpha Omega Trial	68.5	3.33	PASE	Volume	GDS-15 (continuous)	Symptoms
Roh et al, (2015)	Korea	0/1	Change in outcome	6647 (55.5)	Participants from Living Profiles of Older People Survey	69.8	3.00	Two questions on PA frequency and duration	Time	GDS-15 ≥ 8	Screening level
Rubin et al, (2005)	USA	1/0	Change in exposure and outcome	1550 (68.0)	Participants from Diabetes Prevention Program	25+	3.20	Modifiable Activity Questionnaire	MET.hours	BDI ≥ 11	Screening level
Ruiz-Estigarribia et al, (2019)	Spain	0/1	Baseline exposure	14908 (59.5)	Participants from the SUN project	36.7	18.00	Questions on PA duration	MET.hours	SR doctor diagnosis or use of antidepressants	SR diagnosis
Sanchez-Villegas et al, (2008)	Spain	0/1	Baseline exposure	10381 (54.0)	Participants from SUN cohort study	42.1	6.00	Questions on PA duration	MET.hours	SR of physician diagnosis of depression	SR diagnosis
Sin et al, (2016)	USA	1/1	Change in outcome	667 (17.7)	Participants from Heart and Soul Study	66.1	5.00	PA over past month scored on 5-point scale	Frequency	PHQ-9 (continuous)	Symptoms
Smith et al, (2010)	Spain	2/2	Baseline exposure	3196 (0)	Participants from Honolulu-Asia Aging Study	71-93	8.00	One question on distance walked daily	Distance	CES-D ≥ 9	Screening level
Stewart et al, (1994)	USA	0/3	Change in exposure and outcome	736 (59.4)	Participants from Medical Outcomes Study	56.1	2.00	Questions on PA type, frequency, duration, and intensity	Time	MOSFWBP-C (continuous)	Symptoms
Strawbridge et al, (2002)	USA	2/1	Baseline exposure	1947 (56.0)	Participants from Alameda County Study	63.0	5.00	Questions on PA type and frequency	Frequency	DSM diagnosis	Clinical diagnosis
ten Have et al, (2011)	Netherlands	0/2	Baseline exposure	4796 (49.4)	Participants from NEMESIS	18-64	3.00	One question on exercise duration	Time	CIDI computerized version 1.1	Clinical diagnosis
Tsai et al, (2013)	Taiwan	0/2	Baseline exposure	2145 (46.2)	Participants from TLSA	53+	8.00	Questions on frequency, duration, and intensity of PA	Frequency	CES-D-10 ≥ 10	Screening level
Uebelacker et al, (2013)	USA	1/1	Baseline	91912	Participants from	50-79	3.00	Four questions on	MET.hours	CES-D-6 ≥ 5	Screening

			exposure	(100)	WHI-OS			exercise frequency, duration, type, and program duration		level
Uemura et al, (2017)	Japan	2/2	Change in outcome	3106 (48.0)	Participants from Obu Study of Health Promotion for the Elderly	71.5	12.25	Questions on engagement in PA of different intensities	Frequency	GDS-15 ≥ 6	Screening level
Van Gool et al, (2003)	Netherlands	1/0	Change in exposure and outcome	1104 (54.6)	Participants from LASA	69.9	6.00	LASA PA questionnaire	Time	CES-D ≥ 16	Screening level
van Uffelen et al, (2013)	Australia	0/1	Change in outcome	8950 (100)	Participants from ALSWH	50-55	9.00	Questions on PA frequency and duration	METminutes	CES-D-10 ≥ 10	Screening level
Veronese et al, (2017)	UK	1/1	Baseline exposure Change in outcome	4077 (52.5)	Participants from ELSA	70.8	2.00	Questions on frequency of moderate and vigorous PA	Frequency	CES-D ≥ 16	Screening level
Visser et al, (2018)	Netherlands	0/1	Change in exposure and outcome	3107 (55.0)	Participants from LASA	70.0	3.00	LASA PA questionnaire	Volume	CES-D (continuous)	Symptoms
Wang et al, (2011)	Taiwan	0/1	Baseline exposure	197 (46.2)	Community-dwelling older Taiwanese adults	72.5	2.00	Questions on exercise frequency, duration, and type.	Volume	GDS (continuous)	Symptoms
Watts et al, (2018)	USA	0/1	Change in exposure and outcome	594 (59.4)	Participants from the Midwestern Alzheimer's Disease Center Clinical Cohort	72.7	2.00	Rapid Assessment of Physical Activity	Per 1-unit increase	GDS-15 (continuous)	Symptoms
Weyerer, (1992)	Germany	0/2	Baseline exposure	1528 (55.2)	Random sample from three Upper Bavarian communities	47.5	5.00	One question on exercise frequency	Frequency	Clinical interview – ICD diagnosis	Clinical diagnosis
Wise et al, (2006)	USA	12/12	Baseline exposure	35224 (100)	Participants from Black Women's Health Study	21-69	2.00	One question on exercise duration	Volume	CES-D ≥ 16	Screening level
Xie et al, (2019)	China	1/1	Change in outcome	292 (57.0)	Freshmen at Anhui Medical University	19.3	0.50	Single question on PA frequency	Frequency	CES-D (continuous)	Symptoms
Yang et al, (2014)	Finland	0/10	Change in outcome Change in exposure and	1955 (57.5)	Participants from Young Finns Study	31.7	6.00	Questions on PA frequency, intensity, and duration	Volume	Modified BDI (continuous)	Symptoms

			outcome								
Yoshida et al., (2015)	Japan	2/2	Change in exposure	680 (57.2)	Elderly Japanese	72.7	3.00	Question on PA frequency	Frequency	GDS-15 ≥ 6	Screening level
Zhang et al, (2018)	Germany	2/2	Baseline exposure	1196 (100)	Participants from Dresden Predictor Study	18-25	1.42	Single question on PA frequency	Frequency	DSM-IV Major Depression	Clinical diagnosis

Abbreviations: ActiFE=Activity and Function in the Elderly in Ulm; ALSWH=Australian Longitudinal Study on Women's Health; BDI=Beck Depression Inventory; C-SURF=Cohort Study On Substance Use Risk Factors; CARDIA=Coronary Artery Risk Development in Young Adults Study; CePAWHS=Central Pennsylvania Women's Health Study; CES-D=Centre for Epidemiological Studies Depression scale; CIDI=Composite International Diagnostic Interview; CWHS=California Work and Health Survey; DSM=Statistical Manual of Mental Disorders; ELSA=English Longitudinal Study of Ageing; FINE=Finland, Italy, and the Netherlands Elderly; GDS=Geriatric Depression Scale-short form; GHQ=General Health Questionnaire; GLTEQ=Godin Leisure Time Exercise Questionnaire; HADS-D=Hospital Anxiety and Depression Scale-Depression subscale; HUNT=Nord-Trondelag Health Survey; IDS=Inventory of Depressive Symptoms; IPAQ=International Physical Activity Questionnaire; JAGES=Japan Gerontological Evaluation Study; KLoSA=Korean Longitudinal Study of Aging; KLOSCAD=Korean Longitudinal Study on Cognitive Aging and Dementia; LASA=Longitudinal Aging Study Amsterdam; MADRS=Montgomery–Asberg Depression Rating Scale; MDI=Major Depression Inventory; MET=Metabolic Equivalent; MINI-K=Korean Mini-International Neuropsychiatric Interview; MOSFWBP-C=Medical Outcomes Study Functioning and Well-being Profile Core Subset; NEMESIS=Netherlands Mental Health Survey and Incidence Study; NESDA=Netherlands Study of Depression and Anxiety; NHANES=National Health and Nutrition Examination Survey; NILSLSA=National Institute for Longevity Sciences Longitudinal Study of Aging; NLSAA=Nottingham Longitudinal Study of Activity and Ageing; NR=Not reported; NYCNAMES=The New York City Neighborhood and Mental Health in the Elderly Study II; PA=Physical Activity; PASE=Physical Activity Scale for the Elderly; PHQ=Patient Health Questionnaire; PPMI=Parkinson's Progression Markers Initiative; PPSW=Prospective Population Study of Women; SAD=Symptoms of Anxiety and Depression; SAGE=Study on Global AGEing and Adult Health; SHARE=Survey of Health, Aging and Retirement in Europe; SHIP=Study of Health in Pomerania; SHLSE=Survey of Health and Living Status of the Elderly; SMASH=Study on Musculoskeletal disorders, Absenteeism, Stress, and Health; SR=Self Report; SUN=Seguimiento Universidad de Navarra; SWAN=Study of Women's Health Across the Nation; TILDA=The Irish Longitudinal Study on Aging; TLSA=Taiwan Longitudinal Study on Aging; WHI-OS=Women's Health Initiative Observational Study; YEDD=Yeoncheon Elderly Depression and Dementia; ZSDS=Zung Self-rating Depression Scale