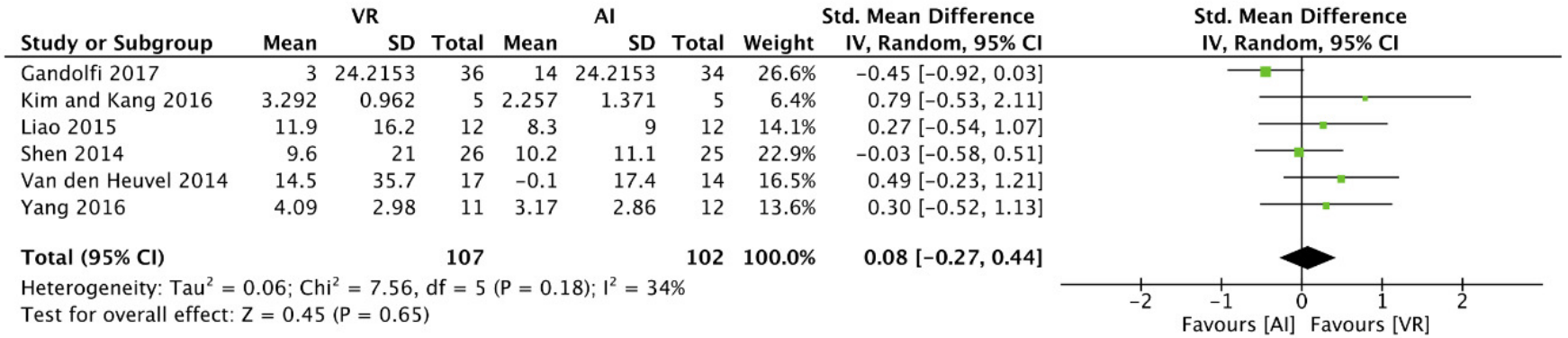
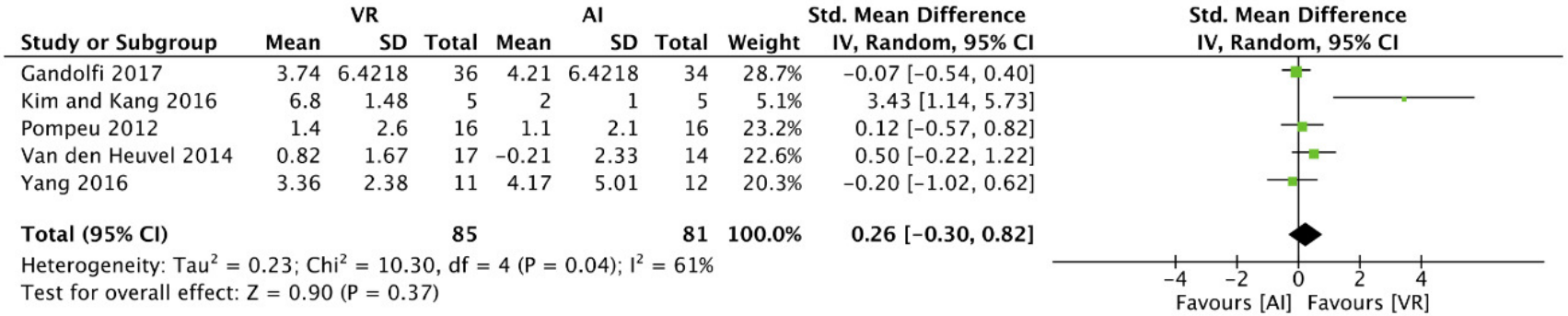
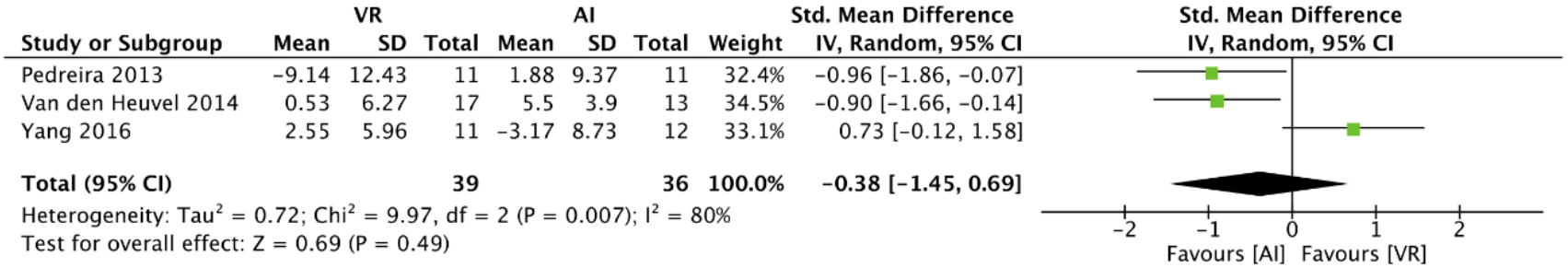
Supplementary Fig. 1 Forest plot of six studies comparing the effects of VR training and active intervention on gait speed.



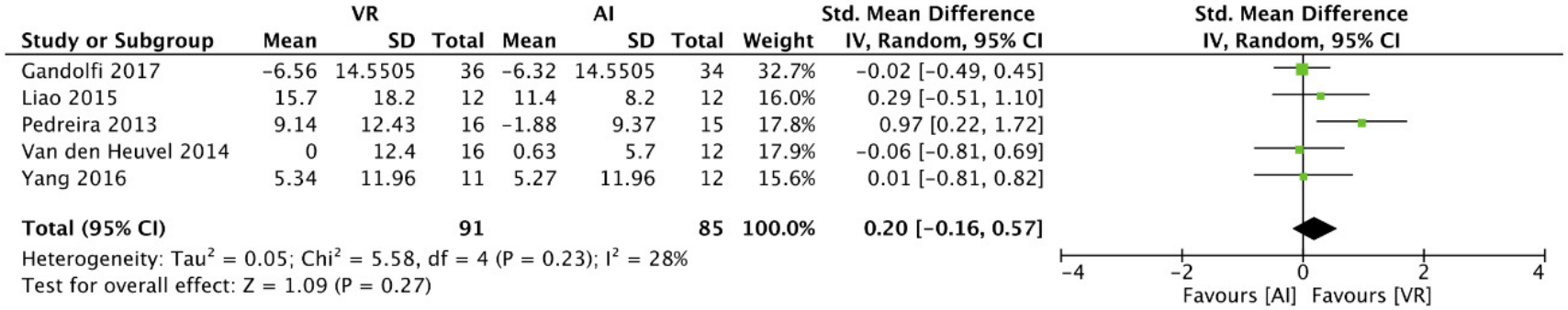
Supplementary Fig. 2 Forest plot of five studies comparing the effects of VR training and active intervention on balance.



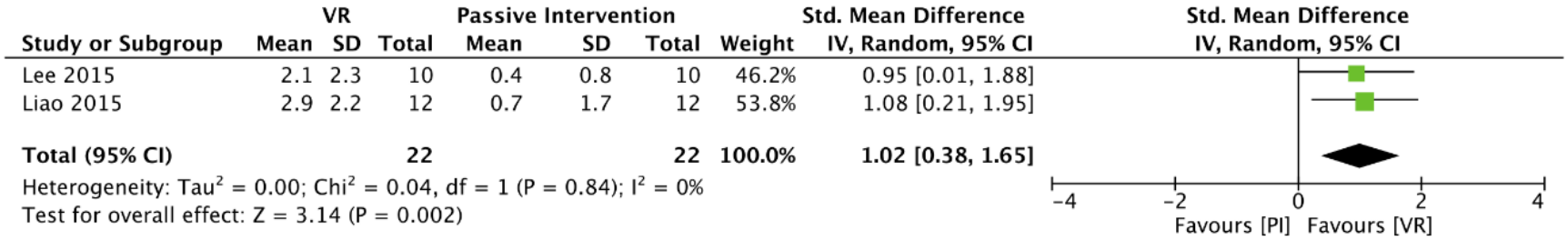
Supplementary Fig. 3 Forest plot of three studies comparing the effects of VR training and active intervention on motor function.



Supplementary Fig. 4 Forest plot of five studies comparing the effects of VR training and active intervention on quality of life.



Supplementary Fig. 5 Forest plot of five studies comparing the effects of VR training and passive intervention on balance.



Supplementary Table 1. Characteristics of systematic review studies.

Studies	n	Mean ±SD age (yrs)	Mean ±SD PD duration (yrs)	Hoehn Yahr scale	PD drugs	M: F	Source of VR	Primary and secondary outcomes
Arias et al (2012)	PD (n=10)	69.9 ±11.2	N/A	N/A	N/A	5: 5	HMD, Vuzix iWear VR920 glasses and motion tracker.	Detect differences in tapping frequency between groups, check the reliability of using VR.
	Young control (n=12)	24.3 ±4.9				7: 5		
	Older control (n=12)	66.6 ±10.1				5: 7		
Badarny et al (2014)	VR (n=20)	71.25	5.28	2.5-4	Y	N/A	Belt-mounted unit, housing motion sensors and digital processing components, and delivered by a micro-display.	Gait speed and stride length.
Cipresso et al (2014)	PD-NC (n=15)	69.0 ±8.1	N/A	N/A	N/A	9: 6	NeuroVR.	Investigate the differences in executive function (EF) and VMET scores between groups.
	PD-MCI (n=15)	68.1 ±9.4				8: 7		
	Control (n=15)	61.7 ±5.2				6: 9		
de Melo et al (2018)	VR=12	60.3 ±9.3	N/A	1.4 ±0.5	Y	11: 1	Xbox Kinect.	Primary: 6MWT (gait speed and endurance) Secondary: physical fitness (SpO_2, Heart rate, blood pressure, BORG).
	Treadmill=13	61.0 ±10.7	N/A	1.5 ±0.7	Y	12: 1		
	Conventional training=12	65.6 ±13.0	N/A	2.1 ±0.9	Y	5: 7		
Tremblay et al (2012)	PD=11	61.9 ±11.0	8.5 ±3.6	N/A	Y	6: 5	Wii Fit.	Primary: evaluate the effect of VR on balance and functional ability of PD patients. Secondary: compare results with healthy people.
	Control=9	63.5 ±12.0				5: 4		
Galna et al. 2014	Game design: 2	N/A	N/A	N/A	N/A	N/A	Xbox Kinect.	Primary: Assess game design Secondary: assess game feasibility.
	Feasibility: 9	68.2 ±8.3	N/A	1.8 ±0.7	Y	3: 6		
Gandolfi et al. 2017	VR=36	67.5 ±7.2	6.2 ±3.8	2.5	Y		Tele-Wii protocol, remote physiotherapist, Wii + balance board, laptop with skype.	Primary: BBS. Secondary: DGI, 10MWT, MCID, PDQ8, satisfaction of patients and comparison of costs between groups.

	SIBT=34	69.8 ±9.4	7.5 ±3.9	2.5	Y			
Herz et al. 2013	VR=20	66.7 ±7.2	5.5 ±4.3	2	Y	13: 7	Nintendo Wii.	Primary: change in NEADL Secondary: quality of life and motor function: changes in the UPDRS, the 9-hole peg test, the Purdue Pegboard Test, a timed tapping task, TUG, HAMD, and the PDQ-39.
Holmes et al. 2013	VR=11	66.6 ±5.9	8.1±3.8	2.3 ±0.4	Y	7: 4	Nintendo Wii.	Balance Centre of Pressure Length (COPL)
Kim et al, 2017	PD-VR=11	65 ±7	N/A	N/A	Y	3: 8	Oculus Rift DK2.	Primary: measure adverse effects of using VR. Secondary: measure levels of arousal.
	Elderly control=11	66 ±3				3: 8		
	Young control=11	28 ±7				5: 6		
Kim & Kang 2016	VR=5	76.2 ±3.9	N/A	2.2 ±0.4	N/A	2: 3	IREX.	Primary: Balance and gait speed Secondary: falls efficacy.
	Control=5	78.4 ±5.8	N/A	2.4 ±0.5	N/A	3: 2		
Lee et al. 2015	VR=10	68.4 ±2.9	N/A	N/A	N/A	5: 5	Nintendo Wii.	Primary: the effect of VR on balance (BBS). Secondary: the effect of VR on activities of daily living and depression (MBI and BDI).
	Control=10	70.1 ±3.3	N/A	N/A	N/A	5: 5		
Liao et al, 2015	VR=12	67.3 ±7.1	7.9 ±2.7	2.0 ±0.7	Y	6: 6	Nintendo Wii Fit.	Primary: gait speed, stride length, obstacle clearance and dynamic balance. Secondary: Sensory organisation test (SOT), timed up and go test, Falls efficacy and PDQ39.
	Traditional exercise=12	65.1 ±6.7	6.9 ±2.8	2.0 ±0.8	Y	6: 6		
	Control=12	64.6 ±8.6	6.4 ±3.0	1.9 ±0.8	Y	5: 7		
Loureiro et al, 2012	VR=6	65 ±13	N/A	2-3	N	N/A	Nintendo Wii Fit.	Primary: Motor skills (Borg scale, BBS, TUG, functional reach tests) Secondary: quality of life (Nottingham scale).
Ma et al, 2011	VR=17	64.8 ±8.5	5.3 ±4.4	2.06 ±0.24	Y	8: 9	OpenGL.	Reaching with favoured hand to grab. real stationery and moving balls down a ramp
	Control=16	68.1 ±7.4	5.2 ±3.4	2.2 ±0.4	Y	10: 6		
Mirelman et al, 2011	VR=20	67.1 ± 6.5	9.8 ±5.6	2.2 ±0.4	Y	14: 6	LED lights on shoes, treadmill and screen.	Primary: gait speed, endurance testing, stride length, obstacle negotiation. Secondary: cognitive and clinical measures (UPDRS and PDQ39)

Messier et al. 2007	PD=8	71.1 ±5.5	10.25 ±3.06	2.6 ±0.2	N	6: 2	SGI Octane.	Accuracy in horizontal and vertical dimension
	Healthy young=10	27	N/A	N/A	Y			
	Healthy elderly=10	68.5	N/A	N/A	Y			
Palacios-Navaro et al, 2015	VR=7	66.8 ±3.5	N/A	N/A	Y		Xbox Kinect.	10MWT
Pedreira et al, 2013	VR=22	61.1 ± 8.2	8.6 ±4.6	2.5 ±0.6	N/A	15: 7	Nintendo Wii.	Quality of life (PDQ39)
	physiotherapy=22	66.2 ±8.5	7.3 ±6.6	2.5 ±0.6	N/A	16: 6		
Pompeu et al. 2012	VR=16	60-85	N/A	N/A	Y	N/A	Nintendo Wii.	Primary: independent performance of daily tasks (UPDRS II). Secondary: Dynamic balance (BBS), static balance (unipedal stance test), cognitive performance (Montreal cognitive assessment).
Pompeu et al, 2014	VR=7	72 ±9	N/A	2.1 ±0.6	Y	6: 1	Xbox Kinect.	Primary: feasibility and safety of using VR (game scores and adverse events). Secondary: clinical outcomes (BESTest, DGI, 6MWT and PDQ39).
Severiano et al. 2018	VR=16	58.7 ±18.7	5.1 ±3.2	N/A	N/A	N/A	Nintendo Wii.	Primary: dizziness handicap index (DHI), BBS and SF36 after 20 VR sessions. Secondary: improvement of scores in each game after 20 VR sessions.
Shen & Mak 2014	VR=22	63.3 ±8.0	8.1 ±4.3	2.4 ±0.5	Y	13: 9	Computerised dancing system + SMART EquiTest balance master.	Primary: balance confidence (ABC). Secondary: balance and gait performance (LOS and SLS.
	Physiotherapy=23	65.3 ±8.5	6.6 ±4.0	2.5 ±0.5	Y	12: 11		
van den Heuvel et al. 2014	VR=17	66.3 ±6.4	9.0 (range: 4.0-13.3)	2.5 (range: 2-3)	Y	12: 5	commercially available interactive dynamic balance exercises (Motek Med- ical, Amsterdam, the Netherlands).	Primary: FRT. Secondary: balance and gait (BBS, SLS, 10MWT). Health status and participation (UPDRS, FES, PDQ39, HADS, MFI).
	Physiotherapy=16	68.8 ±9.7	8.8 (range: 2.5-11.5)	2.5 (range: 2-3)	Y	8: 8		
Yang et al, 2016	VR=11	72.5 ±8.4	9.4 ±3.6	3 (range: N/A)	Y	4: 7	VR balance board.	Primary: BBS. Secondary: DGI, TUG, PDQ39, UPDRS.
	Control=12	75.4 ±6.3	8.3 ±4.1	3 (range: N/A)	Y	5:7		
Yen et al. 2011	VR=14	70.4 ±6.5	6.0 ±2.9	2.6 ±0.5	Y	12:2	VR balance board.	SOT, auditory arithmetic subtraction task.

	Physiotherapy=14	70.1 ±6.9	6.1 ±3.3	2.4 ±0.5	Y	12:2		
	Control=14	71.6 ±5.8	7.8 ±4.2	2.6 ±0.4	Y	9:5		
(Zettergren et al. 2011)	VR=1	69	N/A	N/A	N/A	1	Nintendo Wii Fit.	Primary: measure effect of VR on: gait, TUG, BBS and GDS. Secondary: measure scores of VR games.

Abbreviations: HMD; Head mounted display. VR; virtual reality. PD-NC; Parkinson's disease normal cognition. PD-MCI; Parkinson's disease mild cognitive impairment. VMET; Virtual multiple errands test. 6MWT; 6-minute walk test. SpO_2; peripheral oxygen saturation. CBM; community balance and mobility assessment. TUG; timed up and go test. ABC; activities balance and confidence scale. STST; sitting to standing test. POMA; performance orientated mobility assessment. SIBT; sensory balance integration training. BBS; Berg balance scale. DGI; Dynamic gate index. 10MWT; 10-meter walk test. MCID; minimal clinically important difference. PDQ8; Parkinson's disease questionnaire 8. NEADL; Nottingham Extended Activities of Daily Living Test. UPDRS; unified Parkinson's disease rating scale. HAMD; Hamilton depression scale. PDQ39; Parkinson's disease questionnaire 39. WBDD-EO; weight bearing distribution difference – with eyes open. ML; Mediolateral. MBI; Modified Barthel index. BDI; Beck Depression Inventory. BESTest; Balance evaluation system test. LOS; limit of stability test. SLS; single leg stance test. FRT; functional reach test. HADS; hospital anxiety and depression. MFI; multidimensional fatigue inventory. GDS; geriatric depression scale. N/A; not available.