

	D1	D2	D3	D4	D5	Overall			
de Matos et al. (2009)	+	!	+	+	!	!	+	Low risk	
Mosti et al. (2013)	!	+	+	+	!	!	!	Some concerns	
Liu-Ambrose et al. (2004)	+	+	+	+	!	!	-	High risk	
Harding et al. (2020)	+	+	+	+	+	+			
Kemmler et al. (2020)	+	+	+	+	+	+	D1	Randomisation process	
Kistler-Fischbacher et al. (2022)	+	+	+	+	+	+	D2	Deviations from the intended interventions	
Watson et al. (2018)	+	+	+	+	+	+	D3	Missing outcome data	
Borba-Pinheiro et al. (2016)	!	+	+	+	!	!	D4	Measurement of the outcome	
Basat et al. (2013)	!	+	+	+	!	!	D5	Selection of the reported result	
Watson et al. (2018)	+	+	+	+	+	+			
Kistler-Fischbacher et al. (2021)	+	+	+	+	+	+			
Hettchen et al. (2021)	+	+	+	+	+	+			
Watson et al. (2019)	+	+	+	+	+	+			
Moreira et al. (2014)	!	+	+	+	!	!			
Sen et al. (2020)	!	+	+	+	!	!			
Multanen et al. (2014)	+	+	+	+	+	+			
Multanen et al. (2017)	+	+	+	+	+	+			
Madureira et al. (2010)	+	+	+	+	+	+			
Miko et al. (2017)	+	+	+	+	!	!			
Smulders et al. (2010)	+	+	+	+	!	!			
Evstigneeva et al. (2016)	+	+	+	+	+	+			
Papaioannou et al. (2003)	+	+	+	+	!	!			
Carter et al. (2001)	+	!	+	+	+	!			
Liu-Ambrose et al. (2004)	+	+	+	+	!	!			
Chien, et al. (2000)	!	+	+	+	+	!			
Roghani et al. (2013)	!	+	+	+	+	!			
Wen et al. (2017)	!	+	+	+	+	!			
Jee et al. (2008)	!	+	+	+	!	!			
Choi et al. (2011)	!	+	+	+	!	!			
Kim et al. (2019)	!	+	+	+	+	!			
Feng et al. (2021)	!	+	+	+	+	!			
Wayne et al. (2012)	+	+	+	+	+	+			
Young et al. (2007)	!	+	+	+	!	!			
Jung et al. (2017)	!	+	+	+	!	!			

Supplementary Fig. 1. The result of risk of bias assessment: each risk of bias item for included studies. The 21 studies used adequate methods for allocation sequence generation, and all these applied adequate methods of allocation concealment. The other 13 studies did not provide sufficient information on allocation concealment, so the risk of bias remains unclear. The risk of deviations from the intended interventions was low in 32 studies and unclear in 2 studies. The risk of bias due to missing outcome data was low in all studies. In all studies, outcome measures were appropriate, and the data were not likely to have been influenced by knowledge of the intervention received. The selection bias risk of the reported results was low in 18 studies, unclear in 15 studies, and high in 1 study.