Modified COSMIN criteria used for Risk of Bias Assessment: Validity

Design Requirements	Excellent	Good	Fair	Poor
1. Was the percentage of missing device data given?	Percentage of missing data described-	Percentage of missing data NOT described- just total number of participants included in		
	groups included in the analyses	no indication of how many including in the		
	provided; relative to total	analyses at all (i.e. only the number of		
	number of participants in the	participants in the study is reported, no		
	study	indication how many were included in the		
		devices were missing)		
2. Was there a description of how	Described how missing data	Not described but it can be deduced how	Not clear how missing	
missing data were handled?	were handled- describe	missing	items were handled- no	
	and how they dealt with the	how missing data was dealt with, but can be	table show missing data	
	missing data statistically (i.e.	deduced from table that less data was	but no explanation	
	only data with both devices, or	included in analysis for each group than		
	all available data was included in the analyses)	total number of participants in study		
3. Was the sample size	Adequate sample size	Good sample size	Moderate sample size	Small sample size
included in the analysis adequate?	(≥100)	(50-99)	(30-49)	(<30)
4. Can the criterion used or	Criterion used can be	No evidence provided, but assumable that	Unclear whether the	Criterion used can NOT be
reasonable 'gold standard'?	standard' (evidence provided)	adequate 'gold standard'	considered an adequate	'gold standard' (i.e. self-
			'gold standard'	reported time in activity is
				not a valid reference
5 Ware there any other important	No other important		Other minor	criterion) Other important
flaws in the design or statistical	methodological flaws		methodological flaws	methodological flaws
methods of the study?	ineniodological navis		methodological naws	methodological naws
6. For continuous scores: Were	Percent difference AND	Percent difference only		No PD or way to calculate
percent differences or BA plots or	equivalency OR BA Plot			PD, but has other measures
MAPE calculated?	RMSE CV CCC			for accuracy (BA plot, $M\Delta PE/SE$ of means
	NUMBE, CV, CCC			RMSE, CV, CCC)
7. For continuous scores: Was an	ICC calculated and model or	ICC calculated but model or formula of the	Pearson or Spearman	No ICC or Pearson or
intraclass correlation coefficient	formula of the ICC is described	ICC not described or not optimal. Pearson	correlation coefficient	Spearman correlations
(ICC) calculated?		or Spearman correlation coefficient	evidence provided that no	calculated
		systematic change has occurred	systematic change has	
			occurred or WITH	
			evidence that systematic	
	1		change has occurred	

Modified COSMIN criteria used for Risk of Bias Assessment: Reliability

Design Requirements	Excellent	Good	Fair	Poor
1. Was the percentage of missing device data given?	Percentage of missing data described- number of participants in device groups included in the analyses provided; relative to total number of participants in the study	Percentage of missing data NOT described- just total number of participants included in analysis (no individual group numbers), or no indication of how many including in the analyses at all (i.e. only the number of participants in the study is reported, no indication how many were included in the analyses and if any measures for the devices were missing)		
2. Was there a description of how missing data were handled?	Described how missing data were handled- describe explicitly why data is missing and how they dealt with the missing data statistically (i.e. only data with both devices, or all available data was included in the analyses)	Not described but it can be deduced how missing items were handled- do not state explicitly how missing data was dealt with, but can be deduced from table that less data was included in analysis for each group than total number of participants in study	Not clear how missing items were handled- no information, for example table show missing data but no explanation	
3. Was the sample size included in the analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4. Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws
5. For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated