Table S1. Quality assessment of pre-post designs.

Quality Criteria	Bueno et al., 2015 [86]	Haydicky et al., 2015 [90]	Sibalis et al., 2019 [77]	Van de Weijer-Bergsma et al., 2012 [103]	Zhang et al., 2017 [105]	Zylowska et al., 2008 [78]
Was the study question or objective clearly stated?	Y	Y	Y	Y	Y	Y
Were eligibility or selection criteria for the study population pre-specified and clearly described?	Y	Y	N	N	Y	Y
Were participants representative of those who would be eligible for the intervention in the general or clinical population of interest?	Y	Y	Y	Y	Y	Y
Were all eligible participants that met the pre-specified entry criteria enrolled?	Y	NR	NA	NA	NR	N
Was the sample size sufficiently large to provide confidence in the findings?	N	N	N	N	N	N
Was the intervention clearly described and delivered consistently across the study population?	N	Y	N	Y	N	N
Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants?	Y	Y	Y	Y	Y	Y
Were the people assessing outcomes blinded to the participants' interventions?	Y**	N	NR	N	N	N
Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?	Y, N	Y, N	NR, NR	Y, N	Y, N	N, N
Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?	Y, Y	Y, Y	Y, Y	Y, Y	Y, Y	Y, Y
Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention?	N	Y	N	N	N	N
If the intervention was conducted at a group level did the statistical analysis consider the use of individual-level data to determine effects at the group level?	Y	Y	Y	Y	Y	Y
Evaluation G = Good F = Fair P = Poor	G	G	P	F	F	P

Pre-Post Design Studies evaluated with the NHLBI Study Quality Assessment Tool (Y = Yes, N = No, NR = Not reported, NA = Not applicable). In all cases where NR it was not possible to determine from the reported information. ** indicates that some blind measures were collected but these were accompanied by self-report