

Optional additional table (website) Quality assessment based on the Delphi List [36]. Items of the Delphi List were scored “yes”, “no” or “don't know”. Only items that were assessed with “yes” were given a score of 1 point. A total score for overall methodological quality of maximum 9 points was obtained by applying equal weights to all items. “?” denotes unclear, + denotes yes, – denotes no.

Author	Randomization method	Concealment treatment allocation	Similarity groups at baseline	Eligibility criteria specified	Blinding outcome assessor	Blinding care provider	Blinding patient	Outcome: point estimates and measures of variability	Intention to treat analysis	Total score
Wagena [25]	+	+	+	+	+	+	+	+	+	9
Tønnesen [9]	+	?	+	+	?	-	-	+	+	5
Christenhusz [43, 44]	+	?	-	+	?	-	-	+	-	3
Wilson [37]	+	?	+	+	?	-	-	-	+	4
Borglykke [45]	+	-	+	+	?	-	?	+	+	5
Kotz [39]	+	+	+	+	?	?	+	+	+	7
Tashkin [38]	+	?	+	+	?	?	+	+	+	6
Hilberink [41, 42]	+	?	-	+	?	-	-	+	+	4