

**Supplementary table 3** Methodological quality appraisal using the Newcastle-Ottawa Quality Assessment Scale (NOS) for cohort studies

Study	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Assessment of exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design or analysis	Ascertainment of outcome	Long enough follow-up for outcomes to occur	Adequacy of follow-up of cohorts	Overall score
Fries et al. (1994); Wang et al. (2002); Chakravarty et al. (2008)	-	-	-	-	**	*	*	-	4
Lee et al. (2014, 2016a, 2016b)	*	*	-	*	**	*	*	-	7
Loprinzi (2015)	*	*	*	*	**	*	*	*	9
Oja et al. (2017); Stamatakis et al. (2017)	*	*	*	*	**	*	*	*	9
Schnohr et al. (2000, 2013, 2015, 2018)	*	*	-	*	**	*	*	*	8
Wang et al. (2013)	*	*	*	*	**	*	*	*	9

“-” = criteria not met; “\*” = one point for meeting criteria; “\*\*” = two points for meeting criteria; Overall score: 0-3 points = “poor quality”; 4-6 points = “fair quality”; 7-9 points = “good quality”  
 Lee et al (2016a) = reference number 30; Lee et al (2016b) = reference number 50