

Author/year of publication Newcastle-Ottawa scale (NOS)	Acceptable	Interventions with opioid agonist treatment		Intervention without opioid agonist treatment
		Achmad, Y.M. et al (2009)	Morozova, O. et al (2013)	Sánchez, G.C. et al (2012)
<i>Selection</i>				
<i>Representativeness of the exposed</i>	Representative of adult with substance use disorder in the intervention group	Yes	Yes	Yes
<i>Selection of non-exposed</i>	Drawn from some community as exposed	No	No	No
<i>Ascertainment of exposure</i>	Secured records, structured interviews	Yes	Yes	Yes
<i>Demonstration that outcome of interest was no present at start of study</i>	Only people with SUD infected by HCV, HBV, HIV, or TBC	Yes	Yes	Yes
<i>Risk of low quality (low-medium-high)</i>	Low: 0 "No", medium: 1-2 "No", high: ≥ 3 "No"	Medium	Medium	Medium
<i>Comparability</i>				
<i>Study control for age/sex?</i>	Yes	Yes	Yes	No
<i>Study control for at least three additional factors at baseline?</i>	Injecting substance use, homeless, employment status, ethnicity, marital status, education, HIV RNA, CD4 cell counts, HIV-HCV coinfections, psychiatric disorder, current or previous substance use, prior HBV/HCV/HIV/TBC treatment, advanced fibrosis (Metavir stages F3 or F4)	Yes	Yes	No (Compared two different groups)
<i>Risk of low quality (low-medium-high)</i>	Low: 0 "No", medium: 1 "No", high: 2 "No".	Low	Low	High
<i>Outcome</i>				
<i>Assessment of outcome</i>	Independent blind assessment, record linkage	Yes	Yes	Yes
<i>Was follow-up long enough for outcomes to occur</i>	HCV: SVR (12 weeks posttreatment), HIV: at least three months, TBC: treatment completion, HBV: switch to undetectable HBV DNA, HBsAg, or HBeAg	Yes	Yes	Yes
<i>Adequacy of follow up of cohorts</i>	Complete follow up, or follow up above 80 % of participants initiated treatment	No	Yes	Yes
<i>Risk of low quality (low-medium-high)</i>	Low: 0 "No", medium: 1 "No", high: ≥ 2 "No".	Medium	Low	Low
Score (points)*		7	8	6

* The score is graded from 0 to 10 points based on the three main categories "selection", "comparability" and "outcome". The scale identifies threshold score distinguishing between 'good' and 'poor' quality studies

