

## Appendix 2:

### Quality assessment form adapted from the Ottawa-Newcastle scale (NOS) for assessing non-randomised studies

		Yes/No/Unclear
Selection of participants	[1] Was the inclusion/exclusion clearly described? (for example, age, diagnosis status, MUS) [2] Was inclusion/exclusion assessed using valid and reliable measures? (for example, clinical interview to ascertain MUS or standardised questionnaires) [3] Was recruitment strategy clearly described? [4] Did the investigators ensure that the exposed/unexposed group were comparable (for example did they use stratification or matching)	
Adequate description of study population	[1] Was study population well characterised? ➤ Age ➤ Sex ➤ Ethnicity ➤ Suitable definition of MUS	
Validated method for ascertaining exposure	[1] Was the method used to ascertain exposure clearly defined? [2] Was a valid and reliable measure used to ascertain exposure? (For example what measures were used to confirm MUS) ➤ Standardised questionnaires ➤ Clinical interview	
Validated method to confirm outcome	[1] Was a valid and reliable measures used to ascertain outcome? For example ➤ Mean change in health expenditure ➤ Interviews ➤ Questionnaires	
Adequate follow-up period	[1] Was the follow-up long enough for the outcome to occur? [2] Was the follow-up period the same across all groups?	
Adequate follow-up period	[1] Was follow-up adequate enough for the outcome to occur? [2] Was follow-up period the same across groups? [3] Were differences in follow-up adjusted for using statistical techniques?	
Completeness of follow-up (attrition)	[1] Were drop-out rates and reasons for drop-out similar across exposed and unexposed? [2] Were numbers of dropouts/withdrawals documented at each time point?	
Analysis and controls for confounders	[1] Does the study identify and control for confounders or effect modifiers?	
Sample size calculation	[1] Is the sample size adequate [2] Did the study describe how the sample size was calculated? [3] Was the sample size large enough to detect differences in events between groups? Mean change	
Analytical methods appropriate	[1] What kind of analysis done appropriate for the kind of outcome data? For example, ➤ Continuous – Mixed model, ANCOVA ➤ Categorical - Mixed model for categorical outcome ➤ Dichotomous – Logistic regression [2] Was lost to follow-up accounted for in the analysis (e.g. through sensitivity analysis)	