

Appendix 1: Adapted Newcastle-Ottawa Scale (NOS)

Points awarded for cohort selection, comparability of groups and assessment of outcomes to a maximum of 8 points. In this analysis, 6-8 points considering low ROB, 3-5 medium ROB, 0-2 high ROB.	
Selection: (Maximum 3 points)	
Representativeness of the sample:	<p>(1 point) Truly representative of the average population of children / children with CL/P in the target population (all subjects or random sampling) or somewhat representative of the average in the target population (non-random sampling).</p> <p>(0 points) Selected group of users, limited to single specialty (e.g., paediatric dentists etc), single area/city or within a special group (e.g., members of research networks, dental students etc.) or no description of sampling strategy.</p>
Sample size:	<p>(1 point) Justified and satisfactory.</p> <p>(0 points) Not justified or too small to be representative.</p>
Non-respondents:	<p>(1 point) Comparability between respondent and non-respondent characteristics was established, and the response rate was satisfactory.</p> <p>(0 points) The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory or there was no description of the response rate or characteristics of responders and non-responders.</p>
Comparability: (Maximum 2 points)	
The subjects in different outcome groups (e.g. specialty, setting, age) are comparable, based on the study design or analysis. Confounding factors are controlled:	<p>(1 point) The study controls for the most important factor (select one).</p> <p>(1 point) The study control for any additional factor.</p>
Outcome: (Maximum 3 Points)	
Assessment of the outcome:	<p>(2 points) Independent assessment of practitioner records i.e., external assessor.</p> <p>(1 point) Self report by study author(s), inter and/or intra-rater reliability or validity conducted.</p> <p>(0 points) No description of the assessment process.</p>
Statistical test:	<p>(1 point) The reported descriptive statistics to describe the threshold levels for the population are complete with measures of dispersion (e.g., standard deviation, standard error and range) and if statistical tests are used to analyse associations, these are appropriate, and include confidence intervals and the probability levels (p value).</p> <p>(0 points) The statistical tests were not appropriate, not reported, were incomplete or did not include proper measures of dispersion.</p>