Supplementary Material 2: Quality assessment

In Step 1, initial paper reviews and collation of information were carried out. Each co-author read and reviewed eight to ten publications. The authors used the quality assessment tool that was adapted from [51-53]. Hawker et al.[51] proposed nine categories with detailed descriptions of what constitutes a "good" (4 points), "fair (3 points), "poor" (2 points) or "very poor" (1 point) article, allowing a potential maximum score of 36 (without cut-offs for classifying the total rating for the article):

- Abstract and title: *Did the article provide a clear description of the study?*
- Introduction and aims: *Was there a good background section and clear statement of the aims of the research?*
- Method and data: *Is the method appropriate and clearly explained?*
- Sampling: *Was the sampling strategy appropriate to address the aims?*
- Data analysis: *Was the description of the data analysis sufficiently rigorous?*
- Ethics and bias*: Have the ethical issues been addressed and necessary ethical approval has been gained? Has the relationship between researchers and participants been adequately considered?*
- Findings/results: *Is there a clear statement of the findings?*
- Transferability/generalisability: *Are the findings of this study transferable (generalisable) to a wider population?*
- Implications and usefulness: *How important are these findings to policy and practice?*

Lorenc et al. [52] proposed a classification for the quality grading system with cut-offs, which was used for this study: "high quality (A)" (30-36 points); "medium quality (B)" (24-29 points) and "low quality (C)" (9-24 points). Braithwaite [53] addressed the ambiguity, which arises between the "medium quality" and "low quality" of [52], wherein both has 24 score, thus modified "low quality (C) to the classification of 9-23 points, and was employed for this review. The results of the quality assessment were recorded in a MS Excel spreadsheet. Subsequently, the authors derived and recorded information pertains to each study's characteristics as proposed by Petticrew and Roberts [38]. It includes the type of CSC interventions used in the study, study population, study design and primary outcomes. All recorded information was saved for future analysis.

Step 2 involved an internal peer review process. In this step, each author peer reviewed between eight to ten articles, which were reviewed previously by others in Step 1. According to Wehn et al. [50], this step works as a "quality control mechanism" to ensure that the reviews were thorough and that essential aspects or insights of the reviewed approaches had not been missed, and to reduce subjective judgments about the reviewed impact assessment approaches.

In Step 3, the first and second authors cross-checked the peer review results. Any discrepancies between the initial and the latter peer review results were resolved via discussion among the lead authors. The three-step review process used in this study aims to produce an unbiased and complete review of the 38 publications. As a result, a comprehensive review and discussion of the CSC-SLD phenomenon was attained.