Supplementary Material 1. The Newcastle-Ottawa Scale adapted for dosing accuracy studies.

**Selection (maximum five stars):**
1. Representativeness of the study sample:
   a. Truly representative of the target population (all subjects or random sampling) ★★★
   b. Somewhat representative of the target population (non-random sampling) ★★
   c. Selected group.
   d. No description of the sampling strategy.
2. Sample size:
   a. Justified and acceptable ★
   b. Not justified.
3. Prospective data ★
4. Appropriate application of dose determination technique ★

**Comparability of cohorts based on the design or analysis (maximum two stars):**
1. Paired data ★★
2. Study controls adequately for additional factors ★

**Outcome (maximum three stars):**
1. Assessment of the outcome:
   a. Independent blinded assessment ★★
   b. Independent non-blinded assessment ★
   c. Non-blinded assessment ★
   d. No description

If data analysis was not appropriate for method-comparison studies (no assessment of bias, precision (or repeatability if applicable) and accuracy) then consider downgrading one level.

2. Statistical test:
   a. Appropriate, clearly described statistical analysis with confidence intervals or measure of variance ★
   b. The statistical test is not appropriate, not described, or incomplete.

If study did not group age or weight categories appropriately, or alternatively use percentage error analysis or logarithmic transformation, then downgrade to “b.”