# Tables

TABLE 1: Test-retest reliability metrics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Measure | Type | Formula | Pros | Cons |
| Pearson’s *r* | Relative |  | - Ease of calculation and interpretation  - Scale independent | - Inability to detect systematic error  - Limited to two time points  - Not a measure agreement (accuracy)  - Sample size dependent |
| Intraclass correlations (*ICC*) | Relative |  | - Scale independent  - Multiple types for various scenarios | - Researcher degrees of freedom in choosing the appropriate one  - Not a measure agreement (accuracy) |
| Spearman’s *⍴* | Relative |  | - Ease of calculation and interpretation  - Scale independent | - Inability to detect systematic error  - Limited to two time points  - Not a measure agreement (accuracy) |
| Kendall’s *𝜏* |  |  | - Scale independent | - Inability to detect systematic error  - Limited to two time points  - Not a measure agreement (accuracy) |
| Standard error of measurement (SEM) | Absolute |  | - Contextualizes each measurement | - Scale dependent |
| Bland-Altman Limits of Agreement (LoA) | Absolute | , | - Appropriate for methods comparison  - Contextualizes each measurement | - Limited to two time points  - Biased depending on degrees of freedom (sample size)  - Stringency in detecting meaningful change  - Scale dependent |
| Coefficient of variation | Absolute |  | - Contextualizes each measurement | - Scale dependent |
| T-test, ANOVA |  |  | - Captures systematic change | - Comparison of means only (not individual differences) |