

Greenwood, K.M. & De Nardis, R. (2000). An assessment of the reliability of measurements made using the Melbourne Protocol and the Hanoun Multi-Cervical Unit. Melbourne Whiplash Centre (Manuscript in preparation).

Summary of Findings

The **reliability** of a measurement refers to “the consistency, the reproducibility and the repeatability of the instrument or measurement procedure” (Richman, Makrides & Prince, 1980).

The Reliability Trial

To assess the reliability of measures made using The Melbourne Protocol and the Hanoun Multi-Cervical Unit, a trial was designed in which 26 individuals (who did not have ailments involving the neck) were assessed by three therapists on two occasions each. The trial allowed assessment of inter-observer and intra-observer reliability.

Results:

Inter-Tester Reliability

The consistency of a measurement technique when used by different clinicians over time.

- **Systematic Difference between Therapists**
Results indicate a good degree of agreement between therapists. All averages reported were within 3.3 degrees for ROM measurements and 0.8 lbs for strength measurements.
- **Order of Testing Effects**
There were no systematic differences between the first, second and third measurements. Results indicate that there are no major “warm-up” or familiarisation of technique changes in value and further indicate that the pre-measurement trials conducted in the protocol are sufficient to rule out these effects.
- **Relationship Between the Therapists’ Scores –Correlations**
Correlation coefficients are high (.747 to .949 [approaching 1.0]) indicating good inter-observer reliability.
- **Relationship Between Therapists’ Scores –ICCs**
Intra-Class correlation coefficients are high (.767 to .930 [approaching 1.0]) indicating good inter-observer reliability.
- **Standard Error of Measurement**
SEM’s are low (1.56 to 4.10) indicating good inter-therapist reliability.

Intra-Tester Reliability

The consistency of a measurement technique when used by the same clinician over time.

- **Systematic Changes Over Time**
No systematic differences were identified in scores over time.
- **Relationship Between the Therapists’ Scores – Test-Retest Correlations**
The majority of the correlation coefficients are high (.667 to .895 [approaching 1.0]) indicating good test-retest reliability. ROM extension scores were lower (.529 to .747) indicating some attention is required for this particular measure.
- **Test-Retest Reliability of Therapists’ Scores –ICCs**
The majority of the ICC’s are high (.654 to .879 [approaching 1.0]) indicating good test-retest reliability. Again ROM extension was lower (.531 to .742).
- **Standard Error of Measurement**
SEM’s are low (1.54 to 5.73) indicating good test-retest reliability.
- **Minimum Detectable Change – Test-Retest**
The same therapist over a one week period can reliably detect changes of around 10 degrees in ROM and around 5lbs in strength.