Reliability & Variability

I. Statistical Expressions – What They Mean

A. Mean (**X**)
   1. The average score (sum of scores / number of scores)
      a) Example: 10 students, sum of all grades = 920, **X** = 92.0

B. Correlation Coefficient (**r**)
   1. Commonly referred to as “correlation”
   2. Expresses the relationship between two variables (i.e. height, weight)
   3. Measured on a scale of -1 to +1.
      a) a **–r** value indicates variables move in the opposite direction (i.e., as weight increases, height decreases)
      b) **+r** = as one variable increases, so does the other.
      aa) example: a height-and-weight correlation of r = -.95 would indicate a corresponding height increase to weight decrease
   4. Correlations are measured in terms of strength of the relationship
      a) .90 to 1 = very strong
      b) .80 to .89 = strong
      c) .60 to .79 = moderate
      d) .40 to .59 = moderate to weak
      e) .20 to .39 = weak
      f) .00 to .19 = very weak (no relationship)

C. Standard Deviation (**S.D.**)
   1. The average variance ($\pm$) associated with a mean score
   2. Usually expressed following a mean score
      a) example: mean = 92.0 $\pm$ 6.5

D. Standard Error of the Estimate (**SEE**)
   1. The variation (+) associated with the comparison of two mean variances (S.D.’s)
   2. Often expressed as a percent of the mean of two sets of scores [(SEE/X)* 100]

II. How these statistics are related to Strength Testing

A. Reliability (**r**)
   1. Strength scores must be repeatable (consistent) from test to test.
   2. An r value of .98 between two strength test scores would indicate that a very strong relationship existed between the value of the score on Test 1 and the value of the score on Test 2 (i.e. the test scores are repeatable)

B. Variability (S.D., SEE)
   1. Variability refers to the precision or “closeness” of a group of test scores.
   2. S.D. is used to express the group variability of Test 1, and the group variability of Test 2.
   3. SEE is used to express how close the variability of Test 1 was to the variability of Test 2 (comparison of variances)
   4. A strength test may be repeatable, yet still contain a large amount of variability (e.g.

--- Continued

isokinetic testing)

5. Many factors may affect the variability of strength measurement
    a) amount of sleep
    b) time of day
    c) dietary intake
    d) psychological state
    e) condition of limbs, joints, and muscles
    f) testing equipment
    g) testing design, protocol and procedures (standardization).
    h) knowledge and skill of tester
6. the acceptable variability for most physiological testing is <20%

C. Validity

1. Strength scores may be repeatable, but not necessarily valid
    a) calibration of equipment (i.e., Smidt et al., Spine, October 1989)
    b) limitations associated with the testing technique (isotonic, isokinetic, isometric)

III. Lumbar Extension Machine Test/Retest Reliability & Validity: Trends

A. Mean Values

1. Decrease when moving from flexion to extension (normal curve is descending, from flexion to extension)
2. Increase when moving from Day 1 to Day 2, then plateau, indicating a learning effect associated with the testing
    a) familiarization with testing equipment, environment, and testing demands = less hesitancy on Day 2
    b) neurological adaptation (synchronizing of neuromuscular system)

B. Correlations

1. Decrease when moving from flexion to extension, indicating less consistency
    a) fatigue from multiple joint angle test protocol
    b) difficulty of stabilizing pelvis in extended positions
2. Increase when moving from Day 1 to Day 2 (see A-2 above)

C. Variability

1. Increases when moving from flexion to extension
    a) fatigue from multiple joint angle test protocol
    b) difficulty of stabilizing pelvis in extended positions
2. Decreases when moving from Day 1 to Day 22 (see A-2 above)
3. Overall variability for lumbar extension is 10-15% for healthy normal individuals, and 15-20% for the clinical population. This variation should be considered when determining whether an acute test/retest comparison is acceptable

Reliability and Variability of Isometric Lumbar Extension Strength Testing (N=136)*

	Degrees of Lumbar Flexion						
	0	12	24	36	48	60	72
D1T1 vs D1T2							
r	0.78	0.87	0.93	0.94	0.95	0.96	0.95
Mean	161.1	207.6	233.0	249.5	268.4	291.3	311.3
SEE	46.9	45.2	37.6	27.2	39.1	38.9	46.0
SEE(% mean)	29.1	21.7	16.1	10.9	14.5	13.3	14.8
D1T1 vs D2T1							
r	0.73	0.84	0.91	0.92	0.95	0.94	0.95
Mean	168.2	211.8	239.8	257.0	275.9	295.7	316.6
SEE	54.2	51.5	44.4	46.0	41.2	48.6	47.3
SEE(% mean)	32.2	24.3	18.5	17.9	14.9	16.4	14.9
D2T1 vs D2T2							
r	0.94	0.94	0.95	0.97	0.97	0.98	0.98
Mean	172.4	216.0	241.7	259.0	276.5	297.8	323.7
SEE	29.0	34.0	34.4	29.5	32.6	28.7	30.9
SEE(% mean)	16.8	15.7	14.2	11.4	11.8	9.6	9.5
D2T1 vs D3T1							
r	0.81	0.92	0.94	0.96	0.95	0.94	0.97
Mean	175.5	218.5	245.4	262.0	279.2	297.2	322.9
SEE	47.7	37.2	36.4	33.0	40.9	46.4	36.4
SEE(% mean)	27.1	17.0	14.8	12.6	14.6	15.6	11.2

*Mean and SEE values are N•m. To convert N•m to Ft./lbs., multiply the N•m value by 0.7375.

Data from Graves et al., Quantitative assessment of full range-of-motion isometric lumbar tension strength. *Spine* 15(4), 1990.