A Cost-Effective Alternative to Hand-Held Dynamometers for Shoulder Scaption Strength Measurement in Clinical Practice

Benjamin J Branch[1], Arjun Verma, Robert Walter Rutz, Daniel Hayward, Drayton Daily, Hayden McBride, Patrick Allan Massey, Robert Walter Rutz
[1]Orthopaedic Surgery

INTRODUCTION: Manual muscle testing (MMT) is commonly used to assess shoulder strength in the clinic, but it lacks objectivity and has poor interobserver reliability. Although hand-held dynamometers (HHDs) provide more accurate strength assessment, the cost of these devices limits their accessibility in underserved clinical settings. This study evaluates whether inexpensive digital and analog luggage scales can serve as reliable alternatives to conventional HHDs in measuring shoulder scaption strength.

METHODS: After IRB approval, participants were screened for past shoulder injuries or surgeries before undergoing three shoulder scaption strength tests on each arm. Scaption strength was measured following the method used in the Constant-Murley score assessment. A digital luggage scale, an analog luggage scale, and a standard hand-held dynamometer were used to perform the strength tests. Each subject performed a total of six strength assessments in a randomized device order generated by an online random number generator. The observers were blinded to each other's measurements throughout the testing process to assess interobserver reliability. Devices displayed an output in kilograms, which was recorded for each trial. Inter-rater reliability was determined by Intraclass Correlation Coefficient. The Shapiro-Wilk test was used to determine if the data was normally distributed. Spearman's test was used to determine the correlation between devices, and a Bland-Altman plot was made to determine inter-device agreement.

RESULTS: A total of 24 subjects underwent shoulder scaption testing on both arms, yielding 48 measures. The mean values for shoulder scaption strength for the hand-held dynamometer, digital luggage scale, and analog luggage scale were 9.1kg, 9.2kg, and 7.5 kg, respectively. Intraclass correlation coefficients showed an inter-rater reliability of 0.999 using the analog luggage scale ($p<0.001$). Both the digital and analog scales demonstrated a significant correlation with the HHD with R2 values of 0.913 ($p<0.001$) and 0.952 ($p<0.001$) respectively (figures 1 & 2). Bland-Altman plot demonstrated strong agreement between the devices' measurements with 96% of data points for both the digital and analog scales falling between their respective limits of agreement (figure 3 & 4).

DISCUSSION AND CONCLUSION: The digital and analog luggage scales demonstrated similar reliability to the hand-held dynamometer in this study. Their low cost makes them a practical alternative to hand-held dynamometers without sacrificing the reliability of a traditional hand-held dynamometer. The proven reliability of these inexpensive alternatives, may increase access to strength measurements for research and cost restrictive environments.

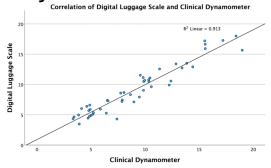

Figure 1: Scatter plot showing a positive correlation between shoulder scaption strength measurements taken by a digital luggage scale or a clinical dynamometer. Spearman's correlation of R=0.934 (p<0.001).

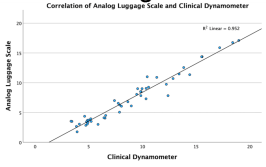

Figure 2: Scatter plot showing a positive correlation between shoulder scaption strength measurements taken by an analog luggage scale or a clinical dynamometer. Spearman's correlation of R=0.957 (p<0.001).

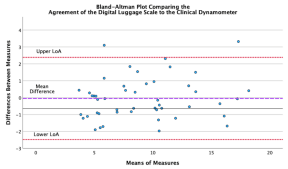

Figure 3: Bland-Altman plot demonstrating agreement of the measures taken by a digital luggage scale and a clinical dynamometer. The mean difference is –0.05. The upper and lower limits of agreement (LoA) are 2.4 and –2.5, respectively. With 96% (46/48) of data points falling between the LoAs, there is a strong agreement between the forms of measure.

LoA: Limit of agreement

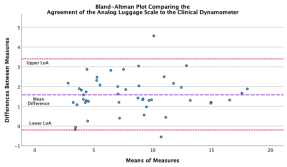

Figure 4: Bland-Altman plot demonstrating agreement of the measures taken by a digital luggage scale and a clinical dynamometer. The mean difference is 1.6. The upper and lower limits of agreement (LoA) are 3.4 and –0.2, respectively. With 96% (46/48) of data points falling between the LoAs, there is a strong agreement between the forms of measure.

LoA: Limit of agreement