Assessment of Inter-Instrument Reliability for Dominant Handgrip Dynamometry and Spirometry

Nnamdi Mgbemena
James Cook University; Australian Institute of Tropical Health and Medicine,
nnamdi.mgbemena@my.jcu.edu.au

Anne Jones
James Cook University; Australian Institute of Tropical Health and Medicine, anne.jones@jcu.edu.au

Anthony Leicht
James Cook University; Australian Institute of Tropical Health and Medicine, anthony.leicht@jcu.edu.au

Recommended Citation

Assessment of Inter-Instrument Reliability for Dominant Handgrip Dynamometry and Spirometry

Abstract

Purpose: The aim of this study was to determine the inter-instrument reliability of different dynamometers and spirometers commonly used in clinical practice. Methods: The study involved 113 healthy volunteers across three facility sites. At each site, dominant handgrip strength (DHGS), and lung function (forced expiratory volume in one second [FEV_1], forced vital capacity [FVC] and peak expiratory flow rate [PEFR]), were compared using a local and reference device. Assessments were randomized with five minutes rest between measurements. Significant differences between devices were assessed using paired t-test while relative reliability between devices was determined via intra-class correlations (ICC). Accuracy index and variability between measurements were assessed using the technical error of measurement (TEM%) and coefficient of variation (CV), respectively. Agreement between devices was determined using the Bland Altman's plot with limits of agreement (LOA). Results: The local devices recorded significantly (p1 (3.1%-8.4%), FVC (3.1%-13%) compared to the reference devices. Good-excellent correlations (ICC=0.89-0.96), unacceptable CV (5.8-9.9%) and TEM% (6.6-9.9%), and large mean biases (3-9kg) and LOA (3-23kg) were identified between the local and reference dynamometers. Excellent correlations (ICC=0.91-0.99), and mostly unacceptable CV and TEM% were identified between the local and reference spirometers for FVC and PEFR. Compared to the reference device, all local spirometers showed unacceptable (-0.134 to -0.536 liters) and acceptable (-0.12 to 0.05 liters/second) mean biases for FVC and PEFR, respectively. Conclusion: Unacceptable inter-instrument reliability was identified between local and reference dynamometers and spirometers for measuring DHGS and all lung function indices, respectively. Across clinical settings, comparing DHGS and lung function between different brands of devices may lead to the reporting of erroneous results with corrective adjustments required for clinical practice.

Author Bio(s)

Nnamdi Mgbemena, MSc.PT, is a PhD candidate (Health) in the College of Healthcare Sciences, James Cook University, Townsville, Australia.

Anne Jones, PhD, is an Associate Professor of Physiotherapy in the College of Healthcare Sciences, James Cook University, Townsville, Australia.

Anthony Leicht is an Associate Professor of Sports and Exercise Sciences in the College of Healthcare Sciences, James Cook University, Townsville, Australia.

Acknowledgements

Assessment of Inter-Instrument Reliability for Dominant Handgrip Dynamometry and Spirometry

Nnamdi Mgbemena
Anne Jones
Anthony Leicht

James Cook University

Australia

ABSTRACT

Purpose: The aim of this study was to determine the inter-instrument reliability of different dynamometers and spirometers commonly used in clinical practice. **Methods:** The study involved 113 healthy volunteers across three facility sites. At each site, dominant handgrip strength (DHGS), and lung function (forced expiratory volume in one second [FEV_1], forced vital capacity [FVC] and peak expiratory flow rate [PEFR]), were compared using a local and reference device. Assessments were randomized with five minutes rest between measurements. Significant differences between devices were assessed using paired t-test while relative reliability between devices was determined via intra-class correlations (ICC). Accuracy index and variability between measurements were assessed using the technical error of measurement (TEM%) and coefficient of variation (CV), respectively. Agreement between devices was determined using the Bland Altman's plot with limits of agreement (LOA). **Results:** The local devices recorded significantly ($p<0.05$) lower mean values for DHGS (7.3-18%), FEV_1 (3.1%-8.4%), FVC (3.1%-13%) compared to the reference devices. Good-excellent correlations (ICC=0.89-0.96), unacceptable CV (5.8-9.9%) and TEM% (6.6-9.9%), and large mean biases (3-9kg) and LOA (3-23kg) were identified between the local and reference dynamometers. Excellent correlations (ICC=0.91-0.99), and mostly unacceptable CV and TEM% were identified between the local and reference spirometers for FVC and PEFR. Compared to the reference device, all local spirometers showed unacceptable (-0.134 to -0.536 liters) and acceptable (-0.12 to 0.05 liters/second) mean biases for FVC and PEFR, respectively. **Conclusion:** Unacceptable inter-instrument reliability was identified between local and reference dynamometers and spirometers for measuring DHGS and all lung function indices, respectively. Across clinical settings, comparing DHGS and lung function between different brands of devices may lead to the reporting of erroneous results with corrective adjustments required for clinical practice.

Keywords: reproducibility of results; hand strength; spirometry; bias; forced expiratory volume.

INTRODUCTION

As part of normal practice, handgrip strength (HGS) and lung function assessments are routinely conducted by allied health professionals to monitor muscular strength and respiratory performance, respectively.[1,2] Reported significant associations between HGS and lung function measures in healthy adults, likely reflects a positive relationship between HGS and respiratory muscle strength, which is a partial determinant of lung function.[3,4] Recently, a systematic review reported on the association between HGS and lung function in both healthy and diseased populations and noted that prior studies utilized inconsistent protocols and a variety of devices for HGS and lung function assessments.[5] Such inconsistencies highlight potential weaknesses for the use of HGS and lung function results by practitioners with clarity needed for effective practice. While consistent protocol adherence can be easily addressed, the consistency between instruments (i.e. inter-instrument reliability) has far greater impact for practitioners and effective practice.[6] Evaluation of inter-instrument reliability within clinical practice has been fundamental to ensure that newer and/or different brands of medical devices assess/monitor patient's progress and prognosis appropriately.[7] For instance, the Jamar hydraulic and the K-Force hand dynamometers were both reported to reliably [intra-class correlation coefficient (ICC) >0.90] assess HGS in healthy and unhealthy adults.[8]

The Jamar hydraulic dynamometer (JHD) is considered as the "gold standard" assessment tool for HGS by the American Society of Hand Therapists (ASHT), however, significant measurement differences were reported among three brands of the JHD.[9] These differences were attributed to inaccuracies with calibration and discrepancies in equipment design with therapists cautioned when comparing different brands and their respective normative values.[9] Prior studies have investigated the inter-instrument reliability of different hand dynamometers using varied measures of reliability and reported conflicting conclusions.[6,8,10-12] For example, Svens et al reported poor instrument reliability between the GripTrack and JHD dynamometers (limits of agreement, LOA, -2.37 kg to 7.87 kg), while Mathiowetz reported good to excellent inter-instrument reliability between the Rolyan and JHD dynamometers (ICC, 0.90-0.97).[10,11] Therefore, poor inter-instrument reliability raises concerns about the use of different hand dynamometer brands during clinical assessments of the same patient (e.g., via multiple clinicians and/or sites) and the need for a greater examination of HGS inter-instrument reliability.

Similarly, heterogeneity exists for the type of spirometers used, and their reliability, for lung function assessment in clinical settings that pose a challenge in monitoring patient's lung function.[13-16] For instance, Milanzi et al in their study involving 49 volunteers, reported a large mean difference (0.24 L) in forced vital capacity (FVC) and forced expiratory volume in one second (FEV$_1$; 0.37 L) between the EasyOne and the Masterscreen pneumotachograph spirometers, and unacceptable reliability between these devices.[16] Others have reported similar poor inter-instrument reliability between spirometer brands with technological differences between spirometers, use of multiple spirometers in-series during testing, and the disease severity of assessed patients as potential contributors.[13-16] Like HGS assessments, spirometry assessments using different brands may be substantially different and unacceptable when utilized in clinical practice.

Despite possible differences between different brands of HGS dynamometers and spirometers, certain situations may necessitate the use of different brands of device in clinical practice or research studies.[6,11,13,16] For instance, in regional areas where multiple sites may assess the same patient/client, the high cost of standard HGS dynamometers and spirometers may not allow each site to possess the same type of device. Subsequently, longitudinal assessments may be conducted using different brands of devices, which may lead to over/underestimation of HGS and lung function values and invalid comparisons. These poor assessments may be further exacerbated by un-standardized HGS and lung function assessment protocols, common among health practitioners.[5] Consequently, the use of standardized assessment protocols with reliable devices may be crucial to clinical practice, particularly within health systems across multiple sites. The aim of this study was to determine the inter-instrument reliability of different HGS dynamometers and spirometers used routinely in clinical practice. It was hypothesized that after adopting standardized assessment guidelines there would be no significant differences between different HGS dynamometers for assessment of muscular strength, and between different spirometers for assessment of lung function.[17] Determination of good inter-instrument reliability of different HGS dynamometers and spirometers would support their use across multiple clinical settings and epidemiological studies.

METHODS
Study Design and Setting
This multi-centered study employed a repeated measures design where each participant completed assessments of HGS and lung function, each using two different devices: a reference and a local device. All assessments were supervised by the same assessor with participants adhering to standardized protocols.[17] Ethics and research governance approvals were obtained from the Human Research Ethics Committee (HREC) at the Townsville University Hospital (HREC/2019/QTHS/53274) with all participants providing written informed consent prior to participation.

Participant Recruitment and Selection Criteria

Participants were a convenient sample of apparently healthy adults, aged between 18 and 80 years, which were recruited from the community through electronic and traditional media sources. A short screening questionnaire was completed by each participant to ensure that none had a history of musculoskeletal disease, hand surgery in the last three months, upper limb deformities, respiratory and/or cardiovascular disease, respiratory tract infection in the last three-weeks, or currently taking medications, which may affect muscle strength or lung function.

Description of the Devices

Four handgrip dynamometers were utilized, and these included:

i. Three Jamar hydraulic dynamometers; [Model 5030J1, Performance Health, China (JHD-PH); Model J00105, Sammons Preston, Bolingbrook, IL (JHD-SP) and Model 5030JI, Patterson Medical, Warrenville, IL (JHD-PM)]. These three similarly designed dynamometers measure a maximum of 90 kilograms (kg) and have five handle positions.[18] Previous studies have reported the JHD as a reliable (test-retest: 0.88 to 0.93; interrater: 0.99) and valid tool for the assessment of HGS.[19,20]

ii. CAMRY electronic dynamometer (Model EH101, Zhongshan Camry Electronic Co., Ltd, China): this is a battery-powered device that records a maximum HGS of 90kg with an ability to be altered to fit different hand sizes during a maximal isometric contraction.[21] Study by Mani et al 2019, has reported excellent test-retest reliability (0.91 to 0.97) of the CAMRY electronic dynamometer in measuring HGS.[22]

Four spirometers were utilised in this study and these included:

i. Vitalograph ALPHA (VA) spirometer (Model 6000, Vitalograph Ltd, Ireland): this is a tabletop device, either mains or battery-powered and uses a disposable bacterial viral filter mouthpiece. The VA spirometer uses the Fleisch-type pneumotachograph for measuring airflow during lung function testing.[23]

ii. EasyOne spirometer (Model 2001, ndd Medical Technologies, Switzerland): this is a battery-powered, hand-held electronic device, which uses a spirette (mouthpiece) with an ultrasonic flow sensor for evaluating an individual's lung function.[24] A prior study by Barr et al, 2009 has reported the EasyOne spirometer as a reliable and valid device in assessing lung function for research and occupational purposes.[25]

iii. CONTEC spirometer (Model SP10, CONTEC Medical systems Ltd, China): this is a battery-powered, hand-held device, which comprises of infrared pair diodes and a microprocessor that measures the airflow only during a forced expiration manoeuvre.[26]

iv. Microlab spirometer (CareFusion, Yorba Linda, CA, USA): this is also a tabletop device, either mains or battery-powered and uses a volume transducer for measuring lung function during inspiration and expiration.[27]

Handgrip dynamometers were calibrated annually while the spirometers were calibrated annually, with daily accuracy checks conducted using a 3-liter syringe before assessments.[1] The JHD-PH and the VA spirometer were selected as the reference devices for all comparisons with local devices, as they were the most commonly used devices at the regional health facilities.

Data Collection Procedures

Assessments were conducted at three different regional health facilities that possessed typical but different HGS and spirometry devices used within clinical practice. At each facility, assessments were completed using the local and reference devices. At facility 1, the local HGS dynamometer and spirometer examined were the JHD-PM and the EasyOne spirometer. For facility 2, the local HGS dynamometer and spirometer examined were the JHD-SP and the Microlab spirometer. At facility 3, the local HGS dynamometer and spirometer examined were the CAMRY dynamometer and the CONTEC spirometer. The order of devices used for assessments, which was only blinded for the participants, was randomized by flipping a coin while HGS assessment was conducted first followed by spirometry, with both assessments separated by five minutes of rest.

Dominant Handgrip Strength Assessment

Participants sat upright on a chair without an armrest, with the shoulders adducted to the side, elbow in 90° flexion, and forearm and wrist in neutral position. Their feet were maintained flat on the ground with hip and knee joints at 90° flexion according to the ASHT guidelines.[2] The Jamar dynamometers were set in the second handle position, as recommended by the ASHT while the Camry dynamometer was adjusted to the third position in order to correspond to the two-centimetre difference between the moveable and stationary arms of the Jamar dynamometers. Standardised instructions were followed by all participants during assessments.[20] Three measurement attempts were undertaken for the dominant hand to achieve a rest phase of at least 15 seconds after every attempt.[2] Each attempt did not last more than six seconds with the highest value recorded as the participant's maximal dominant HGS (DHGS).

Lung Function Assessment
Participants sat upright to ensure the involvement of a forceful and maximal exhalation. A disposable mouthpiece was used for each participant during this assessment and standardised guidelines of the American Thoracic Society (ATS) and European Respiratory Society (ERS) were followed.[1] These guidelines required participants to apply a nose clip, rapidly inhale fully through the mouth followed by maximal exhalation through the mouthpiece until complete lung discharge.[1] Three dynamic lung function indices, FEV_1, FVC and peak expiratory flow rate (PEFR) were assessed. Assessments included three trials, with the highest value for FEV_1, FVC and PEFR utilised in analyses, in accordance with the ATS/ERS repeatability and acceptability criteria.[1]

Statistical Analysis
The demographic characteristics of participants were summarised as mean (standard deviation). Normality of the data was checked via the Shapiro-Wilk test as this was more appropriate for smaller sample sizes (<50).[7] Significant differences between devices were assessed using paired t-test while relative reliability between devices was determined via $ICC_{3,1}$.[7,28] Intra-class correlation coefficients <0.5, 0.5 – 0.75, 0.75 – 0.90 and >0.90 were interpreted as poor, moderate, good and excellent, respectively.[28] Accuracy index and variability between measurements were assessed using the technical error of measurement (TEM) and coefficient of variability (CV), respectively.[29,30] The formulae for these statistical tests are shown here:

TEM $= \sqrt{\sum D^2 / 2N}$
Where D is the difference between repeated measures
N = number of volunteers [29]

CV(%) $= (SD/\bar{x})*100$
Where SD = standard deviation
$\bar{x}$ = Mean [30]

Absolute TEM was converted to relative TEM (%) by dividing the absolute TEM by the mean and then multiplied by 100.[29] Relative TEM and CV of <2% and ≤5%, respectively, were considered acceptable for HGS and lung function measurements as these values indicated good reliability between devices.[29-31] Agreement between the reference and local devices for HGS and lung function measurements was assessed using Bland Altman plots (mean of best trials) and LOA.[7] The level of significance was set as <0.05 for all analyses with analyses conducted using SPSS version 27.0 (IBM Inc, Chicago IL).

Since previous studies have reported inconsistent criteria for determining inter-instrument reliability of HGS dynamometers and spirometers pre-determined criteria were established to confirm reliability between devices.[11-13,16] In the current study, dynamometers were considered reliable for measuring HGS if they fulfilled all of the following four criteria: 1) there was no significant difference between mean values; 2) the ICC value was >0.90; 3) relative TEM and CV values were <2% and ≤5%, respectively; and 4) the LOA were <6.5kg.[28,32] Similarly, spirometers were considered reliable for measuring one or more lung function parameter(s) (FEV_1, FVC and PEFR) if they fulfilled all of the following five criteria: 1) there was no significant difference between mean values; 2) the ICC value was >0.90; 3) relative TEM and CV values were <2% and ≤5%, respectively; 4) the mean difference (bias) was <0.10 L for FEV_1 and FVC, and <0.30 L/sec for PEFR; and 5) the LOA for FEV_1 and FVC did not exceed 0.35 L and 0.50 L, respectively.[28,33,34] Further, when poor reliability was identified between devices, linear regression analyses were conducted to develop equations to correct local device measurements relevant to the reference device.

RESULTS
Demographic Characteristics
Overall, one hundred and thirteen healthy adults (60% males) volunteered across three facility sites for this study. The majority of the participants were young Caucasians, who were predominantly right-handed (Table 1).

Dominant Handgrip Strength and Lung Function Assessments
At each facility site, the local dynamometer produced significantly (p<0.05) lower (7.3-18%) DHGS values compared to the reference device (JHD-PH) (Table 2). Similarly, the FVC values reported for the three local spirometers were significantly lower (3.1%-13%) compared to the reference device (VA spirometer) (Table 2). The EasyOne and CONTEC spirometers were the only local spirometers that produced significantly lower FEV_1 values (3.1%-8.4%) compared to the reference device (VA spirometer) (Table 2). For PEFR, all local spirometers produced similar results to that of the reference device (Table 2).

Table 1: Demographic Characteristics of Participants

	Facility 1 (n=30)	Facility 2 (n=51)	Facility 3 (n=32)
Sex†			
Male/Female	17/13	26/25	25/7
Age (years) ‡	30.9(28.0, 34.5)	38.9(35.6, 42.5)	61.6(56.6, 66.3)
Height (m) ‡	1.76(1.73, 1.79)	1.71(1.69, 1.74)	1.74(1.71, 1.77)
Weight (kg) ‡	77.5(72.4, 83.0)	74.2(69.6, 79.3)	97.8(92.2, 104.4)
BMI (kg/m^2) ‡	24.9(23.7, 26.1)	25.1(24.0, 26.4)	32.4(30.6, 34.7)
Ethnicity†			
Caucasian/African/Asian	22/6/2	29/15/7	31/1/0
Hand Dominance†			
Right/Left	27/3	46/5	31/1

BMI – Body mass index; kg – Kilogram; kg/m^2 – Kilogram/metres^2.
† Data presented as frequency
‡ Data presented as mean (95% Confidence Interval)

Table 2: Comparison of handgrip strength and lung function between reference and local facility devices

		Facility 1 JHD-PH vs. JHD-PM Vitalograph vs. EasyOne	Facility 2 JHD-PH vs. JHD-SP Vitalograph vs. Microlab	Facility 3 JHD-PH vs. CAMRY Vitalograph vs. CONTEC
Mean (SD)	DHGS (kg)	48.9(13.5) vs. 39.7(12.9)*	43.0(12.0) vs. 39.9(11.4)*	46.3(14.3) vs. 38.1(10.6)*
	FEV$_1$ (L)	3.55(0.77) vs. 3.44(0.73)*	3.26(0.74) vs. 3.23(0.76)	3.12(0.82) vs. 2.86(0.72)*
	FVC (L)	4.31(1.05) vs. 4.18(0.94)*	4.06(0.89) vs. 3.82(0.91)*	4.02(1.11) vs. 3.48(1.00)*
	PEFR (L/sec)	8.77(2.33) vs. 8.59(2.38)	8.18(2.04) vs. 8.13(2.08)	8.15(2.15) vs. 8.03(2.15)
ICC (95% CI)	DHGS	0.95 (0.90, 0.98)	0.96 (0.93, 0.98)	0.89 (0.79, 0.95)
	FEV$_1$	0.99 (0.97, 0.99)	0.98 (0.97, 0.99)	0.96 (0.93, 0.98)
	FVC	0.97 (0.94, 0.99)	0.98 (0.96, 0.99)	0.97 (0.93, 0.98)
	PEFR	0.95 (0.90, 0.98)	0.92 (0.86, 0.95)	0.91 (0.82, 0.95)
TEM (%)	DHGS	6.6	5.7	9.9
	FEV$_1$	1.3	3.1	4.9
	FVC	3.9	3.6	5.1
	PEFR	6.1	7.3	8.1
CV (%)	DHGS	6.7	5.8	9.9
	FEV$_1$	2.7	3.1	5.0
	FVC	3.9	3.6	5.2
	PEFR	6.1	7.3	8.1

FEV$_1$= Forced expiratory volume in one second; FVC = Forced vital capacity; PEFR = Peak expiratory flow rate; DHGS = Dominant handgrip strength; kg = Kilogram; L = Liters; L/sec = Liters/second; * = significance at p<0.05; ICC = Intra-class correlation; CI= Confidence interval; TE = Technical error of measurement; CV = Coefficient of variation; JHD-PH = Jamar Hydraulic Dynamometer-Performance Health; JHD-PM = Jamar Hydraulic Dynamometer-Patterson Medical; JHD-SP = Jamar Hydraulic Dynamometer-Sammons Preston; vs. = versus.

Inter-Instrument Reliability Assessment
Good-excellent correlations (ICC = 0.89-0.96) were identified between the local and reference HGS dynamometers (Table 2). The range of values for the relative TEM (%) and CV (%) between local and reference devices were unacceptable for DHGS (6.6-9.9% and 5.8-9.9%, respectively, Table 2). The mean bias for DHGS ranged between 3 and 9 kg for local devices, with the LOA indicating that HGS measurements using the local devices could vary up to 13-23kg from the reference device and unacceptable (Figure 1).

Excellent correlations (ICC = 0.91-0.99) were identified between the local and reference spirometers for all lung function indices in this study. The range of values for the relative TEM (%) between local and reference devices were mostly unacceptable for FEV$_1$ (1.3-4.9%), FVC (3.6-5.1%), and PEFR (6.1-8.1%) (Table 2). All three local spirometers exhibited an acceptable CV for FEV$_1$ while it ranged between acceptable to unacceptable for FVC and PEFR (Table 2). Only the Microlab spirometer exhibited an acceptable mean bias and LOA for FEV$_1$ while other local devices exhibited unacceptable mean biases and LOA for FEV$_1$ (Figure 2). The

mean biases and LOA between the local and the reference spirometers were unacceptable for FVC (Figure 3), while the mean biases between the local and the reference spirometers were acceptable for PEFR (~0.05-0.18 L/s) (Figure 4).
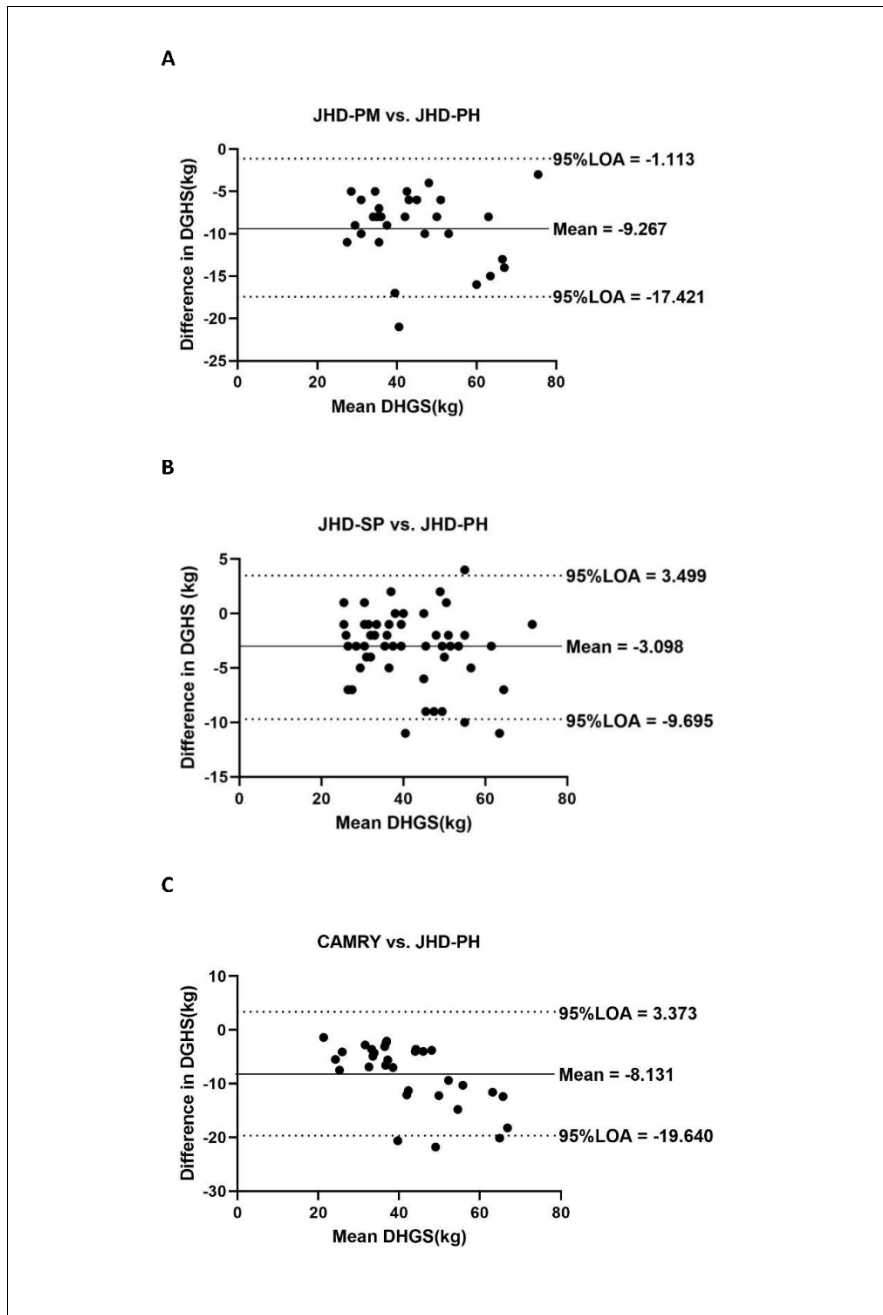

Figure 1. Bland Altman plots of DHGS measurements comparing the JHD-PM (A), JHD-SP (B) and the CAMRY (C) dynamometers to the reference dynamometer (JHD-PH).

JHD-PH = Jamar Hydraulic Dynamometer-Performance Health; JHD-PM = Jamar Hydraulic Dynamometer-Patterson Medical; JHD-SP = Jamar Hydraulic Dynamometer-Sammons Preston; DHGS = Dominant handgrip strength; kg = Kilogram; LOA = Limits of agreement; vs. = versus.
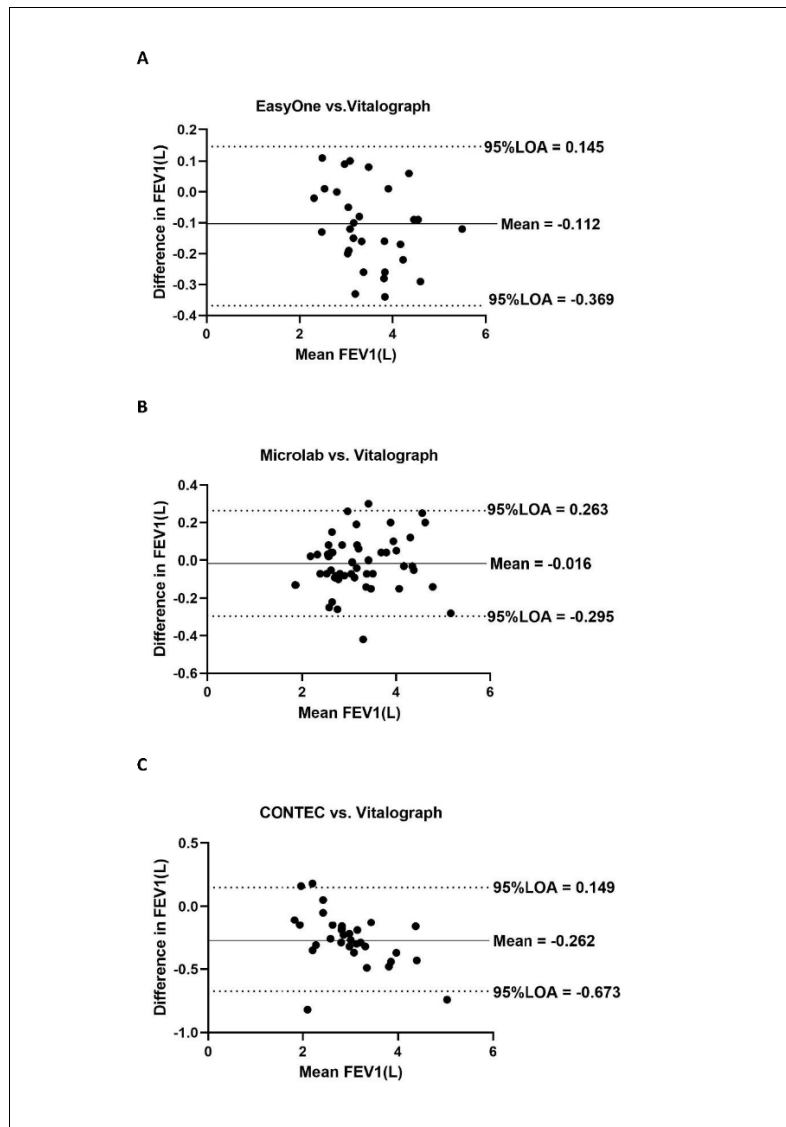
Figure 2 Bland Altman plots of FEV1 measurements comparing the EasyOne (A), Microlab (B) and CONTEC (C) spirometers to the reference spirometer (Vitalograph).

FEV1 = Forced expiratory volume in one second; L = Liters; LOA = Limits of agreement; vs. = versus.
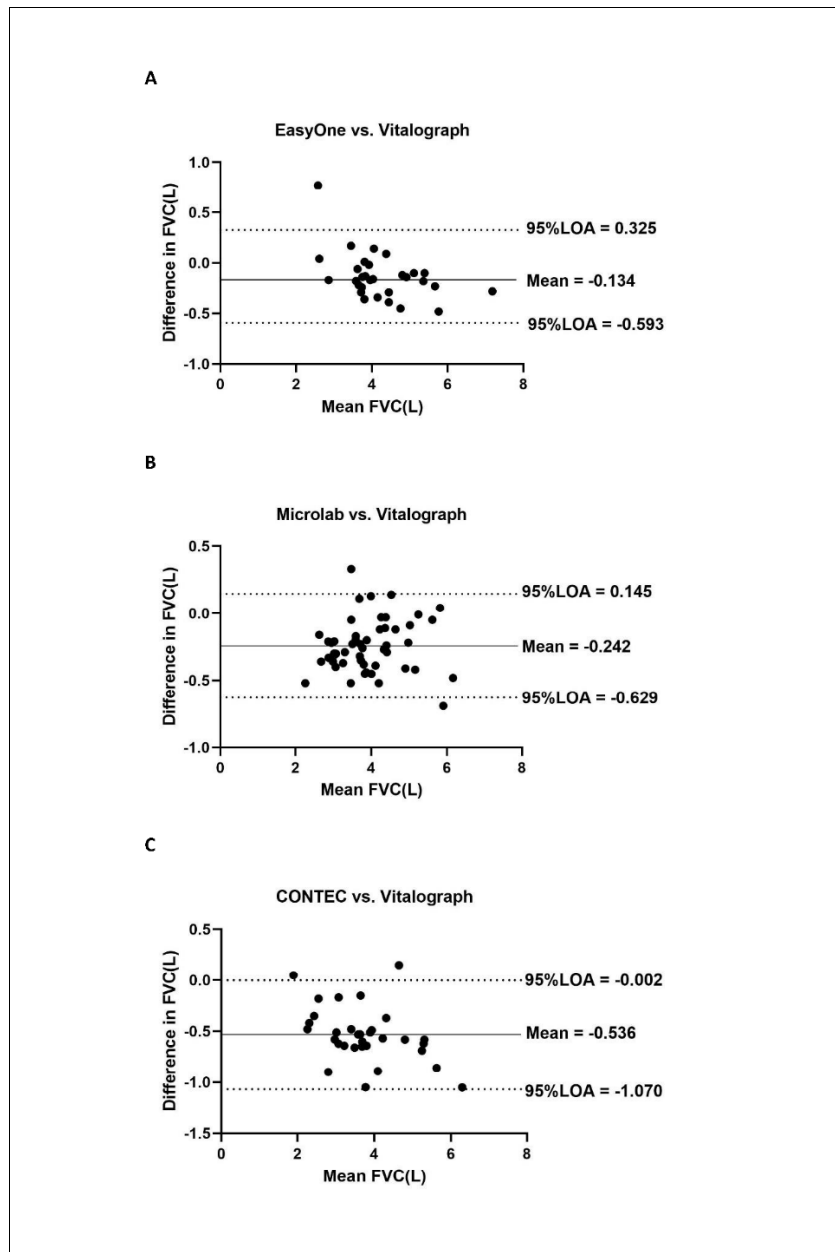
Figure 3. Bland Altman plots of FVC measurements comparing the EasyOne (A), Microlab (B) and CONTEC (C) spirometers to the reference spirometer (Vitalograph).

FVC = Forced vital capacity; L = Liters; LOA = Limits of agreement; vs. = versus.
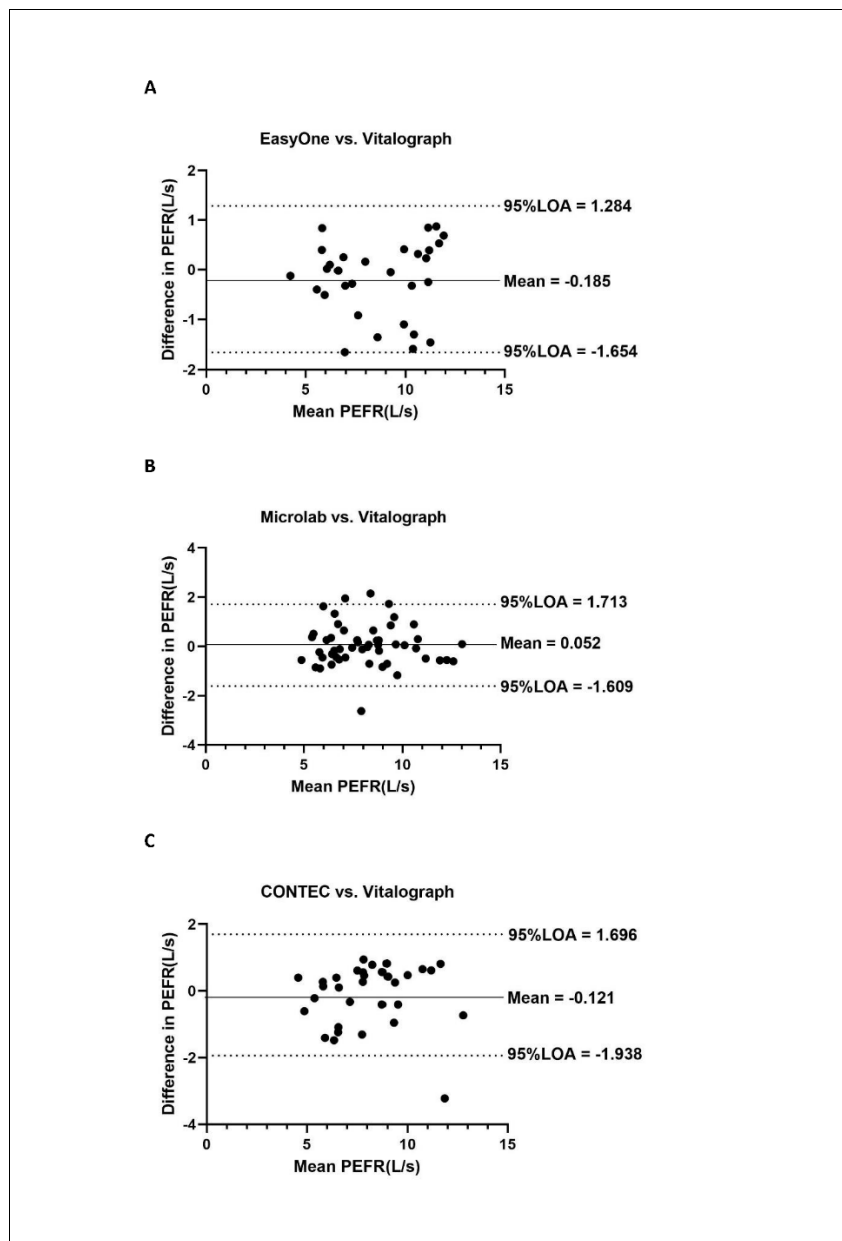
Figure 4. Bland Altman plots of PEFR measurements comparing the EasyOne (A), Microlab (B) and CONTEC (C) spirometers to the reference spirometer (Vitalograph).

PEFR = Peak expiratory flow rate; L/sec= Liters/second; LOA = Limits of agreement; vs. = versus.

To enable reliable comparisons between the local and reference devices, the following linear regression equations were developed to correct HGS, FVC and FEV_1 for local devices:

$$DHGS_{Reference} = 9.506 + 0.994(DHGS_{JHD-PM}) \quad (R^2 = 0.91)$$
$$DHGS_{Reference} = 2.561 + 1.013(DHGS_{JHD-SP}) \quad (R^2 = 0.92)$$
$$DHGS_{Reference} = -1.721 + 1.258(DHGS_{CAMRY}) \quad (R^2 = 0.87)$$
$$FEV_{1Reference} = -0.022 + 1.039(FEV_{1EasyOne}) \quad (R^2 = 0.97)$$
$$FEV_{1Reference} = 0.131 + 0.964(FEV_{1Microlab}) \quad (R^2 = 0.97)$$
$$FEV_{1Reference} = -0.048 + 1.108(FEV_{1CONTEC}) \quad (R^2 = 0.94)$$
$$FVC_{Reference} = -0.246 + 1.091(FVC_{EasyOne}) \quad (R^2 = 0.96)$$
$$FVC_{Reference} = 0.404 + 0.958(FVC_{Microlab}) \quad (R^2 = 0.95)$$
$$FVC_{Reference} = 0.258 + 1.080(FVC_{CONTEC}) \quad (R^2 = 0.95)$$
$$PEFR_{Reference} = 0.794 + 0.929(PEFR_{EasyOne}) \quad (R^2 = 0.90)$$
$$PEFR_{Reference} = 0.502 + 0.932(PEFR_{Microlab}) \quad (R^2 = 0.84)$$
$$PEFR_{Reference} = 0.879 + 0.906(PEFR_{CONTEC}) \quad (R^2 = 0.82)$$

DISCUSSION

This study showed that HGS and most of the lung function assessments with three local devices resulted in significantly lower results compared to their respective reference dynamometer and spirometer. Though excellent correlations were mostly identified between the local and reference dynamometers, the local dynamometers (JHD-PM, JHD-SP and CAMRY) exhibited poor inter-instrument reliability when compared to the reference device (JHD-PH) in measuring DHGS in healthy adults. Compared to the reference device, the three local spirometers exhibited poor inter-instrument reliability in measuring FEV_1, FVC and PEFR. These findings indicated that inter-instrument reliability of HGS and all lung function indices were unacceptable for routinely used dynamometers and spirometers in clinical settings. The use of the current study's regression equations may help clinicians to correct measurements across multiple facilities with different devices.

The observed differences between the three local HGS dynamometers and the reference JHD-PH in the current study corroborates prior studies of poor inter-instrument reliability.[6,9,11,21] While it would be expected that all devices produce similar HGS values due to the similar assessment action (i.e., squeezing device), differences between devices could be attributed to the varied weight of the dynamometer and handle design.[11] In the current study, participants commonly reported that the JHD-PH was lighter and easier to use compared to the other two JHDs (unpublished observation) while the JHD-PM possessed a thicker handle bar compared to the JHD-PH.[11] Therefore, the lower HGS values for the local JHD devices maybe related to the design of the dynamometer. Additionally, the CAMRY dynamometer possessed a different handlebar shape, which may contribute to lower HGS values.[36] The CAMRY dynamometer's handle included clefts for finger placement, which were reported by most participants to cause pain during maximal gripping (unpublished observation). Subsequently, dynamometer design may influence HGS ability and assessments with further work needed to identify the optimal HGS design for reliable assessments. Notably, the current HGS mean differences of greater than 6.5kg between dynamometers, which may be indicative of meaningful clinical changes, indicated poor inter-instrument reliability that clinicians should be aware of when comparing HGS values from different devices across clinical settings.[32] To help clinicians address these concerns across multiple devices, the current study provided correction equations to enable valid clinical comparisons for appropriate therapy within multiple settings.

Previous studies have investigated the inter-instrument reliability between different spirometers and identified poor inter-instrument reliability for most lung function indices.[13-16,34] In the current study, similar PEFR values were identified between spirometers that may be indicative of the ease in assessing larger airway function compared to more vulnerable, smaller airways, which were assessed by FEV_1 and FVC.[37] Conversely, significant differences in FEV_1 and FVC between spirometers were identified that may be explained by the different flow meter systems incorporated within each spirometer. For example, the VA spirometer utilized a Fleisch-type pneumotachograph, which consisted of bundles of small capillary tubes that create a pseudo-laminar flow, and a low-range differential pressure transducer that enables very accurate measurements at all flowrates.[23] In contrast, the CONTEC and Microlab spirometers included a turbine flowmeter with the turbine blades subject to distortion during forced expiratory maneuvers of greater than 5 L/s, causing a lag in the rate of blade rotation and potentially less accurate volumes.[38] Further, the EasyOne spirometer employed a transit-time ultrasound flowmeter with non-uniform laminar and turbulent flow velocities produced during forced expiratory maneuvers resulting in potentially lower and inaccurate volumes.[39] Subsequently, differences in spirometry flow meters between devices may lead to poor inter-instrument reliability for all lung function indices (FEV_1, FVC and PEFR) that clinicians should consider during routine clinical practice across facilities. Therefore, clinicians should use the same spirometer when reassessing patients across multiple sites or employ the correction equations provided in this study to adjust for potential errors in lung function measurements.

The use of inconsistent measures of inter-instrument reliability among prior studies, poses a challenge in comparing lung function and HGS values between respective devices.[6,10,13,16] Considering this shortfall, the current study developed more robust criteria that may be considered by future studies examining different and/or newer brands of dynamometers and spirometers.[40] For instance, similar PEFR values, excellent ICC and an acceptable mean bias were identified between all spirometry devices that indicated good inter-instrument reliability. However, unacceptable relative TEM (%) and CV between spirometry devices were also evident for PEFR values that revealed variation between assessments (i.e., poor inter-instrument reliability). Consequently, the sturdy criteria within the current study may potentially set a standard for future studies and minimize the reporting of misleading conclusions that may have far-reaching clinical implications.

Limitations

The current study utilized robust statistical analyses and criteria for assessing inter-instrument reliability between different hand dynamometers and spirometers.[6,10,13] The use of a trained assessor, accurately calibrated devices, randomized trials and standardized protocols for all assessments minimized potential random errors that may impact the results of the study.[7] However, the study had some limitations. Comparison of a local dynamometer and spirometer to a reference device at each facility site with different populations precluded the assessments of inter-instrument reliability across all four dynamometers or spirometers using a single study population. Secondly, the current study utilized a specific set of common devices used routinely in practice. It remains to be seen whether differences and poor reliability exist for other devices used in clinical practice. Thirdly, inclusion of only healthy participants may limit the application of these findings to unhealthy populations, especially those with respiratory and/or muscular disease conditions who routinely undertake HGS and spirometer assessments.[41] However, the range of HGS and lung function assessments were similar to those seen in unhealthy populations.[42,43] with future studies needed to expand the current findings to unhealthy populations that exhibit lower HGS and lung function measurements.

CONCLUSION

Significant inter-device differences were identified when measuring HGS and lung function via different hand dynamometers and spirometers, respectively. Although good-excellent correlations were observed between the local and their respective reference device, poor inter-instrument reliability defined via robust criteria was identified between these dynamometers and spirometers. The use of different brands of dynamometers and spirometers for assessing and/or monitoring HGS and lung function, respectively, may lead to the reporting of erroneous results with corrective adjustments required for clinical practice.

REFERENCES

1. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of spirometry 2019 update. An official American Thoracic Society and European Respiratory Society technical statement. *Am J Respir Crit Care Med*. 2019;200(8):e70-e88. doi:10.1164/rccm.201908-1590ST. [PMID: 31613151]
2. MacDermid J, Solomon G, Fedorczyk J, Valdes K. *Clinical assessment recommendations: Impairment-based conditions.* 3rd ed. American Society of Hand Therapists; 2015.
3. Shin Hi, Kim D-K, Seo KM, Kang SH, Lee SY, Son S. Relation Between Respiratory Muscle Strength and Skeletal Muscle Mass and Hand Grip Strength in the Healthy Elderly. *Ann Rehabil Med*. 8 2017;41(4):686-692. doi:10.5535/arm.2017.41.4.686. [PMID: 28971054]
4. Mgbemena NC, Aweto HA, Tella BA, Emeto TI, Malau-Aduli BS. Prediction of lung function using handgrip strength in healthy young adults. *Physiological Reports*. 2019;7(1):e13960. doi:10.14814/phy2.13960. [PMID: 30632320]
5. Mgbemena N, Jones A, Leicht AS. Relationship between handgrip strength and lung function in adults: a systematic review. *Physiother Theory Pract*. 2021:1-20. doi:10.1080/09593985.2021.1901323. [PMID: 33870831]
6. King TI, II. Interinstrument reliability of the Jamar electronic dynamometer and pinch gauge compared with the Jamar hydraulic dynamometer and B&L Engineering mechanical pinch gauge. *Am J Occup Ther*. 2013;67(4):480-483. doi:10.5014/ajot.2013.007351. [PMID: 23791323]
7. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 3rd ed. Pearson Education; 2009.
8. Nikodelis T, Savvoulidis S, Athanasakis P, Chalitsios C, Loizidis T. Comparative Study of Validity and Reliability of Two Handgrip Dynamometers: K-Force Grip and Jamar. *Biomechanics*. 2021;1(1):73-82. [PMID: Not Available]
9. Flood-Joy M, Mathiowetz V. Grip-strength measurement: a comparison of three Jamar dynamometers: occupation, participation and health. *The Occupational Therapy Journal of Research*. Jul/Aug 1987;7(4):235-243. [PMID: Not Available]

10.   Mathiowetz V. Comparison of Rolyan and Jamar dynamometers for measuring grip strength. *Occup Ther Int*. 2002;9(3):201-9. doi:10.1002/oti.165.[PMID: 12374997]

11.   Svens B, Lee H. Intra- and inter-instrument reliability of Grip-Strength Measurements: GripTrack™ and Jamar® hand dynamometers. *The British Journal of Hand Therapy*. Jun 2005;10(2):47-55. doi:10.1177/175899830501000202. [PMID: Not Available]

12.   Fike ML, Rousseau E. Measurement of adult hand strength: a comparison of two instruments: occupation, participation and health. *The Occupational Therapy Journal of Research*. Jan 1982;2(1):43-49. [PMID: Not Available]

13.   Berntsen S, Stølevik SB, Mowinckel P, Nystad W, Stensrud T. Lung function monitoring; a randomized agreement study. *Open Respir Med J*. 2016;10:51-57. doi:10.2174/1874306401610010051. [PMID: 27583055]

14.   Caras W, Winter M, Dillard T, Reasor T. Performance comparison of the hand-held microplus portable spirometer and the sensormedics Vmax22 diagnostic spirometer. 1999.

15.   Gerbase MW, Dupuis-Lozeron E, Schindler C, et al. Agreement between spirometers: a challenge in the follow-up of patients and populations? *Respiration*. 2013;85(6):505-514. doi:10.1159/000346649. [PMID:23485575]

16.   Milanzi EB, Koppelman GH, Oldenwening M, et al. Considerations in the use of different spirometers in epidemiological studies. *Environ Health*. Apr 2019;18(1):39. doi:10.1186/s12940-019-0478-2. [PMID: 31023382]

17.   Bialocerkowski AE, Bragge P. Measurement error and reliability testing: Application to rehabilitation. *International Journal of Therapy and Rehabilitation*. 2008;15(10):422-427. doi:10.12968/ijtr.2008.15.10.31210. [PMID: Not Available]

18.   Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age Ageing*. 2011;40(4):423-429. doi:10.1093/ageing/afr051. [PMID: 21624928]

19.   Fess EE. A method for checking Jamar dynamometer calibration. *Journal of Hand Therapy*. 1987;1(1):28-32.

20.   Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *The Journal of Hand Surgery*. 1984;9(2):222-226.

21.   Muñoz G, Millán S. Comparing the Camry dynamometer to the Jamar dynamometer for use in healthy Colombian adults. *Revista Salud Bosque*. Dec 2019;9:21-29. doi:10.18270/rsb.v9i2.2794. [PMID: Not Available]

22.   Mani P, Sethupathy K, Francis AK. Test-retest reliability of electronic hand dynamometer in healthy adults. *International Journal of Advanced Research*. 2019;7(5):325-331.

23.   Vitalograph. Vitalograph alpha model 6000: instructions for use December 2020. Accessed January 26, 2021 https://vitalograph.eu/downloads/view/294

24.   ndd Medical Technologies. EasyGuide Operator's Manual 2009. Accessed January 26, 2021 https://d3cumwfln92soa.cloudfront.net/5335c56cebc1cf0412e5a51dca9dbaf9d99b83b1/easyoneplusmanual.pdf

25.   Barr R, Stemple K, Mesia-Vela S, et al. Reproducibility and Validity of a Handheld Spirometer. *Respiratory care*. 2008;53:433-41.

26.   CONTEC Medical Systems. SP-10 pocket spirometer with bluetooth: user manual January 2019. Accessed January 26, 2021 https://www.gimaitaly.com/DocumentiGIMA/Manuali/EN/M33535EN.pdf

27.   CareFusion. CareFusion Microlab operating manual December 2009. Accessed January 26, 2021 https://micromedical.wms.co.uk/documents/product-manuals/MicroLab8.pdf

28.   Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg Am*. Mar 1984;9(2):222-226. doi:https://doi.org/10.1016/S0363-5023(84)80146-X. [PMID: 6715829]

29.   Koo TK, Li MY. A guideline of selecting and reporting Intraclass Correlation Coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012. [PMID: 27330520]

30.   Oliveira T, Oliveira G, Ornellas J, Oliveira F. Technical error of measurement in anthropometry (English version). *Revista Brasileira de Medicina do Esporte*. Feb 2005;11:81-85. [PMID: Not Available]

31.   Zady MF. Z-4: Mean, standard deviation and coefficient of variation 2019. Accessed July 19, 2021 https://www.westgard.com/lesson34.htm#coefficient

32.   Sylvester KP, Clayton N, Cliff I, et al. ARTP statement on pulmonary function testing 2020. *BMJ Open Respir Res*. 2020;7(1):e000575. doi:10.1136/bmjresp-2020-000575. [PMID: 32631927]

33.   Bohannon RW. Minimal clinically important difference for grip strength: a systematic review. *J Phys Ther Sci*. Jan 2019;31(1):75-78. doi:10.1589/jpts.31.75. [PMID: 30774209]

34.   Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26(2):319-338. doi:10.1183/09031936.05.00034805. [PMID: 16055882]

35.   Liistro G, Vanwelde C, Vincken W, Vandevoorde J, Verleden G, Buffels J. Technical and functional assessment of 10 office spirometers: A multicenter comparative study. *Chest*. Sep 2006;130(3):657-65. doi:10.1378/chest.130.3.657. [PMID: 16963659]

36.   Amaral JF, Mancini M, Novo JM. Comparison of three hand dynamometers in relation to the accuracy and precision of the measurements. *Rev Bras Fisioter*. 2012;16(3):216-224. doi:10.1590/s1413-35552012000300007. [PMID: 22801514]

37. Verbanck S. Physiological measurement of the small airways. *Respiration*. 2012;84(3):177-188. doi:10.1159/000341742. [PMID: 22948000]

38. Mandal NG. Respirometers including spirometer, pneumotachograph and peak flow meter. *Anaesthesia & Intensive Care Medicine*. Jan 2006;7(1):1-5. doi:https://doi.org/10.1383/anes.2006.7.1.1. [PMID: Not Available]

39. Zhang H, Guo C, Lin J. Effects of velocity profiles on measuring accuracy of transit-time ultrasonic flowmeter. *Applied Sciences*. Apr 2019;9(8):1648. doi:10.3390/app9081648. [PMID: Not Available]

40. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. *PLoS One*. 2012;7(5):e37908-e37908. doi:10.1371/journal.pone.0037908. [PMID: 22662248]

41. Kukull WA, Ganguli M. Generalizability: the trees, the forest, and the low-hanging fruit. *Neurology*. 2012;78(23):1886-1891. doi:10.1212/WNL.0b013e318258f812. [PMID: 22665145]

42. Cichosz SL, Vestergaard ET, Hejlesen O. Muscle grip strength is associated to reduced pulmonary capacity in patients with diabetes. Article. *Prim Care Diabetes*. 2018;12(1):66-70. doi:10.1016/j.pcd.2017.06.007. [PMID: 28705696]

43. Strandkvist VJ, Backman H, Röding J, Stridsman C, Lindberg A. Hand grip strength is associated with forced expiratory volume in 1 second among subjects with COPD: Report from a population-based cohort study. *Int J Chron Obstruct Pulmon Dis*. 2016;11(1):2527-2534. doi:10.2147/COPD.S114154. [PMID: 277785009]