sensors

MDPI

Systematic Review

The Reliability of the Microsoft Kinect and Ambulatory Sensor-Based Motion Tracking Devices to Measure Shoulder Range-of-Motion: A Systematic Review and Meta-Analysis

Peter Beshara [1,2,3,*], David B. Anderson [4], Matthew Pelletier [2,3] and William R. Walsh [2,3]

1 Department of Physiotherapy, Prince of Wales Hospital, Sydney, NSW 2031, Australia
2 Prince of Wales Clinical School, Faculty of Medicine, University of New South Wales, Sydney, NSW 2031, Australia; m.pelletier@unsw.edu.au (M.P.); w.walsh@unsw.edu.au (W.R.W.)
3 Surgical & Orthopaedic Research Laboratories, Prince of Wales Hospital, Sydney, NSW 2031, Australia
4 School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW 2006, Australia; david.anderson1@sydney.edu.au
* Correspondence: Peter.Beshara@health.nsw.gov.au

Abstract: Advancements in motion sensing technology can potentially allow clinicians to make more accurate range-of-motion (ROM) measurements and informed decisions regarding patient management. The aim of this study was to systematically review and appraise the literature on the reliability of the Kinect, inertial sensors, smartphone applications and digital inclinometers/goniometers to measure shoulder ROM. Eleven databases were screened (MEDLINE, EMBASE, EMCARE, CINAHL, SPORTSDiscus, Compendex, IEEE Xplore, Web of Science, Proquest Science and Technology, Scopus, and PubMed). The methodological quality of the studies was assessed using the consensus-based standards for the selection of health Measurement Instruments (COSMIN) checklist. Reliability assessment used intra-class correlation coefficients (ICCs) and the criteria from Swinkels et al. (2005). Thirty-two studies were included. A total of 24 studies scored "adequate" and 2 scored "very good" for the reliability standards. Only one study scored "very good" and just over half of the studies (18/32) scored "adequate" for the measurement error standards. Good intra-rater reliability (ICC > 0.85) and inter-rater reliability (ICC > 0.80) was demonstrated with the Kinect, smartphone applications and digital inclinometers. Overall, the Kinect and ambulatory sensor-based human motion tracking devices demonstrate moderate–good levels of intra- and inter-rater reliability to measure shoulder ROM. Future reliability studies should focus on improving study design with larger sample sizes and recommended time intervals between repeated measurements.

Keywords: Kinect; reliability; range of motion; inertial sensor; meta-analysis

check for updates

Citation: Beshara, P.; Anderson, D.B.; Pelletier, M.; Walsh, W.R. The Reliability of the Microsoft Kinect and Ambulatory Sensor-Based Motion Tracking Devices to Measure Shoulder Range-of-Motion: A Systematic Review and Meta-Analysis. *Sensors* **2021**, *21*, 8186. https://doi.org/10.3390/s21248186

Academic Editor: Gregorij Kurillo

Received: 7 November 2021
Accepted: 30 November 2021
Published: 8 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The clinical examination of individuals with shoulder pathology routinely involves the measurement of range-of-motion (ROM) to diagnose, evaluate treatment, and assess disease progression [1–3]. The shoulder complex involves the coordination of the acromioclavicular, glenohumeral and scapulothoracic joints, to allow motion in three biomechanical planes, specifically the sagittal, coronal, and axial planes [4]. Forward flexion and elevation occur in the sagittal plane; abduction and adduction occur in the coronal plane; and internal and external rotation occur along the long axis of the humerus [5].

The shoulder joint's complex multiplanar motion presents a challenge for clinicians to accurately measure ROM and upper limb kinematics [6,7]. Prior attempts to implement a global coordinate system to describe shoulder movement and define arm positions in space [8] have failed to gain clinical consensus due to practical difficulties. The biomechanical complexity of the shoulder is demonstrated by the synergy of movements necessary for a person to perform activities of daily living. Activities such as reaching for a high

shelf or hair washing requires a combination of flexion and adduction. Similarly, reaching behind the back pocket requires a combination of internal rotation, extension, and adduction. Although many models have been proposed in literature, it nevertheless remains difficult to determine the contribution of individual components of glenohumeral joint and scapulothoracic joint motions. Therefore, the reliability of any tools used for ROM measurement is important for clinicians to make informed decisions regarding patient management [9].

According to the American Academy of Orthopaedic Surgeons (AAOS), normal active ROM of the shoulder is 180° for flexion and abduction, and 90° for external rotation [10]. However, a number of age and health-related variables exist that can influence shoulder ROM, including factors such as gender, work history, and hand-dominance. Studies have demonstrated an overall reduction in ROM across all shoulder movements with age in subjects with no shoulder pathology [11,12]. Gill et al. [13] reported age-related decreases in right active shoulder flexion by 43° in males, and 40.6° in females, and right active shoulder abduction by 39.5° and 36.9°, respectively. Authors also noted a decline in external rotation range, particularly among females. The age-related causes of decreased ROM occur from sarcopenia (loss of muscle mass due to a decrease in Type-II fibres), changes in fat redistribution and slower collagen fibre reproduction, leading to reduced elasticity and shortened ligaments and tendons [14,15].

A goniometer is the most commonly used instrument by clinicians to measure joint position and ROM [16,17]. It is essentially a 360° protractor, comprising a stationary arm, a movable arm, and a fulcrum. When used correctly by a trained clinician, the benefits of goniometry's low cost and portability [18,19] are offset by the disadvantages of low inter-reliability [20,21] and measurement variability between clinicians [22–25]. Furthermore, the two-handed requirement of using a goniometer makes it difficult to stabilise the trunk and scapula, resulting in an increased likelihood of measurement error [26]. Alternatively, inclinometers measure ROM relative to the line of gravity and demonstrate improved inter-rater reliability compared with goniometry when assessing shoulder measurements [27,28]. The drawbacks of inclinometery include higher cost, poorer accessibility, and possible technical errors due to misplacing body landmarks or not sustaining constant pressure with the device during movements.

With the increasing popularity, accessibility, and convenience of smartphones and similar devices [29], the potential exists for these electronics to become a clinician's measuring tool of choice. Smartphones with inbuilt accelerometers and magnetometers can utilise inclinometer or sensor-based applications to calculate shoulder joint angles [30–32]. Similarly, digital inclinometers or goniometers are compact, portable, and lightweight. However, a degree of training is required for the clinician to accurately determine a zero point and limit measurement error [33]. Although more costly than traditional manual methods, digital inclinometers and smartphones eliminate the need for realignment and require only one hand to operate [34]. Additionally, the ability to transmit measurements may decrease transpositional or other manual entry errors.

Further advancements in technology over the last decade have led researchers to adopt hands-free motion sensing input devices to estimate human joint ROM [35]. The Microsoft Kinect sensor was originally developed as an add-on for the Xbox 360 gaming console (Microsoft Corp., Redmond, WA, USA) [36] and has since been modified for application in real-world settings including telehealth [37,38], education [39,40] and kinematic motion analysis [41–43]. The Kinect sensor combines a regular colour camera with a depth camera that comprises an infrared laser projector and infrared camera. The Windows Software Development Kit (SDK) 2.0 has allowed for the creation of applications that utilise the Kinect's gesture recognition capability to implement joint orientation and skeletal tracking for 25 joint positions in the standing or seated positions [44,45]. Given its potential breadth of use, the Kinect is emerging as a promising clinical tool for kinematic analysis by virtue of its function as a markerless system to estimate the 3-D positions of several body joints [46].

Ambulatory sensor-based human motion tracking devices such as inertial measurement units (IMUs) comprise accelerometers, gyroscopes, and magnetometers. IMUs measure linear acceleration and angular velocities, combining data to obtain a 3-D position and orientation of a body [47]. The miniaturisation, wearability, and low cost of IMUs over the last decade has made it a desirable alternative to expensive motion-capturing systems for measuring joint angles [48–50]. Prior studies evaluating the precision of IMUs have reported mean average errors of <5.0° for upper limb measurements [51–53]. However, IMU accuracy varies depending on the amount of ROM a joint can produce [53], type of device [54], and fusion algorithm used [55].

Prior to using any device for ROM assessment in clinical practice, it is important to establish the measurement properties of validity and reliability [56]. Several studies have previously validated the Microsoft Kinect [57–60], IMUs [61–63], digital inclinometers [64–67] and smartphone applications [68–70] against a prescribed "gold standard" for shoulder ROM measurement.

For the purposes of this review, authors examined reliability, reflecting the degree to which measurements are consistent over time and across different observers or raters [71]. The two recognised types of reliability in the literature are: intra-rater reliability—the amount of agreement between repeated measurements of the same joint position or ROM by a single rater, and inter-rater reliability—the amount of agreement between repeated measurements of the same joint position or ROM by multiple raters [72].

Absolute reliability is considered equally important and indicates the amount of variability for repeated measurements between individuals [73,74]. Examples include the standard error of measurement (SEM), coefficient of variation (CV), and Bland and Altman's 95% limits of agreement [74]. Absolute measures of reliability allow clinicians to evaluate the level of measurement error and determine whether any changes in ROM signify a real change in their patients [75,76].

Few studies have summarised or appraised the literature on the reliability of the Microsoft Kinect, IMUs, smartphones, and digital inclinometers for human joint ROM measurement. Previous systematic reviews have focused on applying the Kinect for stroke rehabilitation [77–79], and Parkinson's disease [80–82]. Only one systematic review on reliability was identified, which examined using the Kinect to assess transitional movement and balance [83]. To our knowledge, no systematic appraisals of studies on the intra- and inter-rater reliability of the Kinect and ambulatory sensor-based motion tracking devices for shoulder ROM measurement have been conducted.

Therefore, the aim of this article is to review systematically, and appraise critically, the literature investigating the reliability of the Kinect and ambulatory sensor-based motion tracking devices for measuring shoulder ROM.

The specific study questions for this systematic review were:

1. What is the intra- and inter-rater reliability of using the Microsoft Kinect, inertial sensors, smartphone applications, and digital inclinometers to calculate a joint angle in the shoulder?
2. What are the types of inertial sensors, smartphone applications, and digital inclinometers currently used to calculate a joint angle in the shoulder?
3. What clinical populations are utilising motion-tracking technology to calculate the joint angle in the shoulder?
4. Which anatomical landmarks are used to assist the calculation of joint angle in the shoulder?

2. Materials and Methods

The protocol for this review was devised in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement guidelines [84], published with PROSPERO on the 8 December 2017 (CRD 42017081870).

The search strategy was developed and refined by previous systematic reviews investigating reliability [85,86]. A database search of Medline (via OvidSP), EMBASE (via OvidSP),

EMCARE (via Elsevier), CINAHL (via Ebsco), SPORTSDiscus (via Ebsco), Compendex (via Engineering Village), IEEE Xplore (via IEEE), Web of Science (via Thomson Reuters), Proquest Science and Technology (via Proquest), Scopus (via Elsevier), and Pubmed was initially performed on 30 January 2020 by two independent reviewers (PB, DB). These databases were searched from their earliest records to 2020. An updated search was completed on 17 December 2020. Details of the search strategy are found in Supplementary S1. The reference lists of all included studies were screened manually for additional papers that met the a priori inclusion criteria.

2.1. Inclusion and Exclusion Criteria

Studies were included if they met the following criteria: published in peer-reviewed journals; measured human participants of all ages; used the Microsoft Kinect, inertial sensors, smartphone applications, or digital inclinometers to measure joint ROM of the shoulder joint and assessed the intra- and/or inter-rater reliability of these devices; published in English and had full text available. Case studies, abstracts only and "grey" literature was not included. Studies only investigating validity, scapular or functional shoulder movements were excluded, as the aim of the review was to examine the reliability of specific shoulder joint movements commonly measured in clinical practice.

The titles and abstracts of studies were retrieved using the search strategy (Supplementary S1) and screened independently by two review authors (PB, DB). Full text versions that met the selection criteria were uploaded to an online systematic review program (Covidence) for independent review by both reviewers (PB, DB). Any disagreements on eligibility were initially resolved by discussion between reviewers and resolved by a third reviewer (WRW), if necessary.

2.2. Data Extraction

A standardised, pre-piloted form was used to extract data from the included studies for assessment of study quality and evidence synthesis. The following information was extracted for each study: bibliometric (author, title, year of publication, funding sources); study methods (study design, country, setting, description and number of raters, type of shoulder joint movements; type of movement (active ROM (aROM) or passive (pROM)); number of sessions, session interval, type and description of technology); participants (recruitment source, number of drop outs, sample size, age, gender inclusion criteria); anatomical landmarks, statistical methods (type of reliability), and outcomes (intraclass correlation coefficient (ICC) values).

2.3. Evaluation of Reliability Results

Reliability was assessed using ICCs; an ICC value approaching 1 was indicative of higher reliability. The level of intra- and inter-rater reliability was determined by the criteria identified by Swinkels et al. [87]. Intra-rater reliability was considered good with an ICC > 0.85, moderate with ICCs 0.65–0.85, and poor with an ICC < 0.65. Inter-rater reliability was considered good with an ICC > 0.80, moderate with ICCs 0.60–0.80, and poor with an ICC < 0.60.

2.4. Evaluation of the Methodological Quality of the Studies

The two review authors independently assessed the methodological quality of each included paper using the latest (2020) Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) Risk of Bias tool [88]. The studies were rated against a specific set of criteria, with nine items assessing reliability standards and eight items assessing measurement error standards. To satisfy item seven of the measurement error standards, the study had to report absolute reliability statistics (standard error of measurement (SEM), smallest detectable change (SDC) or Limits of Agreement (LOA)). Each item was graded on a four-point scale as either very good, adequate, doubtful, or inadequate. The worst-score-count method was applied in accordance with the COSMIN protocol; the

overall score was determined by the lowest score awarded for the measurement property, as used in previous studies [89,90].

2.5. Data Analysis

Meta-analyses of relative intra- and inter-rater reliability were performed for studies with outcome measures that reported comparable data. Pooled analysis was completed for maximal aROM and pROM. The right-hand dominant value for the healthy, asymptomatic population was included for analysis. Studies with multiple reliability values were pooled and one overall mean result was reported. If a single study reported values for more than one rater, the mean value was reported. Reflecting clinical practice, any reliability values taken in supine position were included in the pROM analysis, and the standing or sitting positions were included in the aROM analysis.

3. Results

3.1. Flow of Studies

A flowchart of the different stages of the article selection process is outlined in Figure 1. From the 2006 studies identified, 32 studies [91–122] were found to meet the criteria for inclusion. In total, nine studies reported reliability for the Microsoft Kinect; six studies for wearable inertial sensors; seven studies for smartphone/mobile applications; and ten studies for digital inclinometers or goniometers.

3.2. Description of Studies

The characteristics of the included studies are summarised in Table 1. A total of 1117 participants were included in this review, with a mean age ranging from 17.0 to 56.1 years of age. The mean sample size was 35 participants with a considerable range (minimum, 1; maximum 155) and variance (SD, 32.1). Six studies recruited more than 50 participants and five studies recruited fewer than ten participants. In 13 of the studies, there was a higher percentage of women compared to men. Most studies (n = 26) recruited participants who were healthy and asymptomatic. Participants with shoulder pain or pathology were reported in six studies.

A physical therapist (PT) was the most reported type of rater (n = 12 studies). In six studies the rater was a medical practitioner (MP), and in two studies a PT student was the sole primary rater. Thirteen studies did not report the profession of the rater.

The shoulder movements assessed across all studies included flexion, extension, abduction, external rotation, internal rotation, and scaption. A total of 24 studies only assessed aROM; eight studies assessed pROM, and two assessed both. The most common measuring position was standing (n = 10 studies), followed by seated (n = 6 studies) and supine (n = 3 studies). There were twelve studies that used a combination of supine and standing, side-lying, prone or seated positions. Only one study did not report the position used.

The majority of studies (n = 25) reported two sessions; five studies had one session, and two studies involved three sessions. The time interval between assessments varied considerably from 10 s to 7 days. The most common consecutive measurements were on the same day (n = 13) followed by 7 days (n = 5).

3.3. Intra and Inter-Rater Reliability

Results for intra- and inter-rater reliability are shown in Table 2. The last column of Table 2 indicates the level of reliability, grouped by type of device, and includes the shoulder movement assessed.

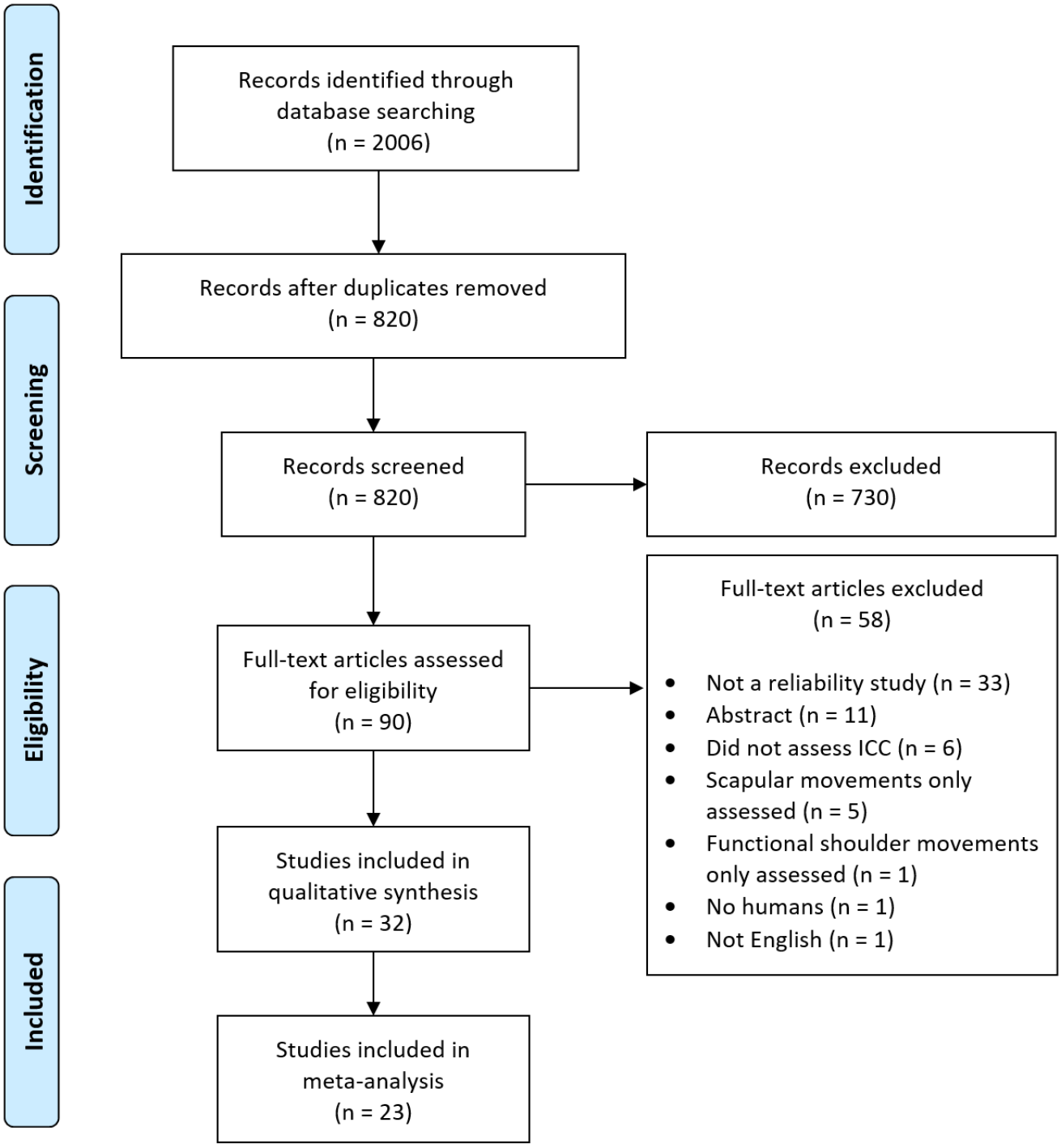

Figure 1. Flow chart of the systematic review process.

Table 1. Characteristics of studies included in this review.

Author	Sample Size (*n*)	Age (yr) Mean (SD)	Males (%)	Inclusion Criteria	Rater (*n*, Profession)	Movement Assessed	Position	Device	Sessions (*n*)	Time Interval
Awan et al., 2002 [91]	56	Not reported	57.1	No history of neurologic disease, arthritis, connective tissue disorder, or shoulder/neck injury or surgery	2 PT MP	Passive • IR • ER	Supine	Digital inclinometer	2	90–120 min
Beshara et al., 2016 [92]	9	36.6 (±13.3)	33.3	No history of neurologic disease, arthritis, connective tissue disorder, or shoulder/neck injury or surgery	1 PT	Active • F • Abd	Standing	Microsoft Kinect (V.2) and Inertial sensors	2	7 days
Bonnechère et al., 2014 [93]	48	26 (±8)	62.5	Healthy adults	1	Active • Abd	Standing	Microsoft Kinect (V.1.5)	2	7 days
Cai et al., 2019 [94]	10	24.6 (±2.8)	100	No upper limb injuries or medication use that would have influenced their upper limb functions	1	Active • F • E • Abd • Add • IR • ER	Standing	Microsoft Kinect (V.2)	2	7 days
Chan et al., 2010 [95]	1	Not reported	100	Healthy, no pathology	2	Active • F • E • RER in 90° Abd	Standing Supine	iPod touch	2	Same day
Chen et al., 2020 [96]	10	Not reported	Not reported	Healthy, aged 20–70 yrs, no discomfort or limited ROM of shoulder in the last year	2 -1 PT -1 MP	Active • F • E • Abd • IR • ER	Standing	Inertial sensor (BoostFix)	1	Same day
Cools et al., 2014 [97]	30	22.1 (1.4)	50	No history of shoulder or neck pain or current participation in overhead sports on a competition level	2	Passive • ER • ER in 90° • Abd • IR in 90° • Abd • IR in forward • F	Sitting Supine	Digital inclinometer	2	10 s

Table 1. *Cont.*

Author	Sample Size (*n*)	Age (yr) Mean (SD)	Males (%)	Inclusion Criteria	Rater (*n*, Profession)	Movement Assessed	Position	Device	Sessions (*n*)	Time Interval
Correll et al., 2018 [98]	42	32.3 (2.1)	71.4	Healthy, 18–75 yrs old, able to easily move between standing and supine positions, able to actively move at least one shoulder into 90° of glenohumeral abduction	2 PT student	Active • F • Abd • ER • IR	Supine	Digital inclinometer (HALO)	2	Same day
Çubukçu et al., 2020 [99]	40	22.1 (±3.1)	55	Healthy volunteers	1 PT	Active • F • E • Abd • ER • IR	Standing	Microsoft Kinect (V.2)	3	3 days
Cuesta-Vargas et al., 2016 [100]	37	56.1 (Healthy) 52.8 (Pathologic)	40.5	Healthy: no shoulder pain, negative NEER/Hawkin's testPathologic: 18–75 yrs old, BMI 18–42	2 PT	Active • Abd	Standing	Inertial sensors (Inertia-Cube 3)-Sampling frequency 1000 Hz Smartphone (Nexus 4) 1280 × 768p resolution	3	2 days
Da Cunha Neto et al., 2018 [101]	10	Notreported	Not reported	Healthy	2	Active • F • E • Abd • Add	Standing	Microsoft Kinect (V.2)	2	Same day
De Baets et al., 2020 [102]	10	54 (±6)	57.1	Diagnosis of adhesive capsulitis in the past 6 months based on criterial described by the American Physical Therapy Association	2	Active • F • E • Abd • Add • IR • ER	Standing Seated	Inertial sensor (MCN Awinda motion capture system)-Sampling frequency 60 Hz	2	2–5 days
de Winter et al., 2004 [103]	155	47	35.5	Shoulder pain, 18–75 yrs, ability to co-operate (no dementia), sufficient knowledge of Dutch language	2 PT	Passive • Abd • ER	Seated Supine	Digital inclinometer (Cybex EDI 320)	1	1 h

Table 1. *Cont.*

Author	Sample Size (*n*)	Age (yr) Mean (SD)	Males (%)	Inclusion Criteria	Rater (*n*, Profession)	Movement Assessed	Position	Device	Sessions (*n*)	Time Interval
Dougherty et al., 2015 [104]	90	23.5 (8.9)	40	18 yrs +, pain free shoulder movement, no history of shoulder pain in preceding 12 months	1 PT	Passive • FGH • F • Abd • GH • Abd • ER in neutral • Abd • ER in 90° • Abd • IR in 90° Abd	Seated Supine	Digital inclinometer	2	7 days
Hawi et al., 2014 [105]	7	Not Reported	Not Reported	Age 18+, free ROM without deficits	1	Active • FE • Abd • Add	Standing	Microsoft Kinect	2	Same day
Huber et al., 2015 [106]	10	22.1 (±0.9)	60	No shoulder pathology, pain-free	1	Active • F to 90° F to max • Abd to 90° • Sagittal F to 90° • Sagittal F to max • ER to max	Standing	Microsoft Kinect	1	Same day
Hwang et al., 2017 [107]	8	36.5 (±13.7)	Not Reported	Using a wheelchair for 1 yr, able to sit upright for at least 4 h/day, over 18 yrs old, use a wheelchair over 40 h/week	1	Active • F • E • Abd • Add	Seated	Microsoft Kinect (V.2)	2	Same day
Kolber et al., 2011 [108]	30	25.9 (3.1)	40	Asymptomatic adults	2 PT	Active • F • Abd • IR • ER	Seated Supine Prone	Digital inclinometer (Acumar)	2	2 days
Kolber et al., 2012 [109]	30	26 (4.2)	30	No cervical spine or upper extremity pain or recent shoulder surgery on dominant arm	2 PT student	Active • Scaption	Seated	Digital inclinometer (Acumar)	2	1 day

Table 1. *Cont.*

Author	Sample Size (*n*)	Age (yr) Mean (SD)	Males (%)	Inclusion Criteria	Rater (*n*, Profession)	Movement Assessed	Position	Device	Sessions (*n*)	Time Interval
Lim et al., 2015 [110]	47	24.9 (±3.5)	59.6	No shoulder injuries or history of musculoskeletal and nervous system damage that could affect ROM, no pain around shoulder no performance of specialized shoulder muscle stretch or exercises or stretching in preceding 6 months	2 PT	Passive • Abd	Supine Side-lying	Smartphone (iPhone 5)	2	2 days
Mejia-Hernandez et al., 2018 [111]	75	46	72	Older than 18 yrs, documented current shoulder diseases	2 MP	Active & Passive • Forward F • Abd • GH • Abd • IR • ER	Seated Supine	Smartphone (iPhone 5)	2	Same day
Milgrom et al., 2016 [112]	5	Not reported	80	Possess ability to self-propel a manual wheelchair, uses a wheelchair for at least 75% of daily activities, ≥18 yrs of age	3 Kinect sensors "individual rater"	Active • F • Abd	Seated	Microsoft Kinect (V.1.8)	2	Same day
Mitchell et al., 2014 [113]	94	Not reported	0	No shoulder pathology	5 -2 PT -3 PT students	Active • ER	Supine	Smartphone (iPhone 4)	2	At least 15 min (<30 min)
Picerno et al., 2015 [114]	45	M: 27 (±8) F: 22 (±3)	55.6	No previous or current shoulder impairment, no involvement in competitive sports at a professional level	1	Active • Abd	Seated	Inertial sensor (FreeSense)-Sampling frequency 200 Hz	2	Same day
Poser et al., 2015 [115]	23	44	39.1	Asymptomatic people who are attending a Pilates gym	3 PT	Active • ER • IR • Abd • Hor • Add	Supine Seated Side-lying	Digital Inclinometer (J-Tech)	2	Days (un-specific)

Table 1. *Cont.*

Author	Sample Size (*n*)	Age (yr) Mean (SD)	Males (%)	Inclusion Criteria	Rater (*n*, Profession)	Movement Assessed	Position	Device	Sessions (*n*)	Time Interval
Ramos et al., 2019 [116]	54	26.3 (6) Healthy 25 (6) Shoulder pain	25.9	Healthy: Not reported Shoulder pain: Symptoms for at least 6 months and positive clinical tests for shoulder impingement	1	Active • F • Scaption	Seated	Mobile application (iPod)	2	7 days
Rigoni et al., 2019 [117]	30	32.8	40	Healthy volunteers	2	Active • F • Abd • ER • IR	Standing	Inertial Sensor (Biokin)	1	Same day
Schiefer et al., 2015 [118]	20	37.4 (±9.9)	70	Healthy subjects without or with known functional deficits, free of musculoskeletal complaints for at least one week before examination	3 MP	Passive • ER • IR	Not reported	Inertial Sensor (CUELA system)	1	1 day
Scibek et al., 2013 [119]	11	21.4 (±1.4)	55.6	Healthy, reporting no history of neck, upper extremity pathology in the last six months	Not reported	Active • F • GH • F • Abd	Seated	Digital inclinometer (Pro 360, Baseline)	2	12–48 h
Shin et al., 2012 [120]	41	52.7 (±17.5)	48.8	Unilateral symptomatic shoulders	3 MP	Active & Passive • Forward F • Abd • ER • ER at 90° • Abd • IR at 90° • Abd	StandingSupine	Smartphone (Galaxy S)	2	Same day
Walker et al., 2016 [121]	17	17 (±3)	47	Healthy, competitive swimmers, at least five swim sessions per week	2 PT	Active • EL • EI • RER • Abd in IR	SupineStanding	Digital inclinometer (Dualer, J-Tech)	2	30 min

Table 1. *Cont.*

Author	Sample Size (*n*)	Age (yr) Mean (SD)	Males (%)	Inclusion Criteria	Rater (*n*, Profession)	Movement Assessed	Position	Device	Sessions (*n*)	Time Interval
Werner et al., 2014 [122]	24	Not reported	37.5	Healthy and symptomatic shoulders, college students	5 -4 MP -1 Medical student	Active • Forward F • Abd • ER at 0° • ER at 90° • IR at 90° • Abd	SupineStanding	Smartphone (iPhone)	2	Same day

Abd = abduction, Add = adduction, ELE = elevation, ER = external rotation, E = extension, ER= external rotation, F = flexion, GH = Glenohumeral, Hor = horizontal, IR = internal rotation, Max = maximum, MP = medical physician, PT = physiotherapist, ROM = range of motion.

Table 2. Intra-rater and Inter-rater reliability (95% CI) for measurement of shoulder range of motion by device and movement direction.

Device	Author	Intra-Rater Reliability	Inter-Rater Reliability	Level of Reliability
Microsoft Kinect				
Shoulder				
Flexion	Da Cuncha Neto et al. (2018) Hawi et al. (2014) Huber et al. (2015) Hwang et al. (2017) Milgrom et al. (2016) Çubukçu et al. (2020) Cai et al. (2019)	ICC 0.97 ICC 0.99 ICC 0.37, 0.85, 0.84, 0.95 ICC 0.96 (0.83–0.98), 0.92 (0.89–0.95) ICC 0.85 ICC 0.93, 0.99, 0.97. 0.96	ICC 0.91 ICC 0.97 (0.84–1.00)	*Good* *Good* *Poor–Good* *Good* *Good* *Moderate* *Good*
Extension	Da Cuncha Neto et al. (2018) Hawi et al. (2014) Hwang et al. (2017) Çubukçu et al. (2020) Cai et al. (2019)	ICC 0.97 ICC 0.99 ICC 0.96 (0.83–0.98), 0.92 (0.89–0.95) ICC 0.62 ICC 0.93, 0.99, 0.97, 0.96	ICC 0.97	*Good* *Good* *Good* *Poor* *Good*
Abduction	Bonnechère et al. (2014) Hawi et al. (2014) Huber et al. (2015) Hwang et al. (2017) Milgrom et al. (2016) Cai et al. (2019)	ICC 0.73 ICC 0.96 ICC 0.76 ICC 0.92 (0.89–0.93), 0.96 (0.86–0.96) ICC 0.70, 0.75, 0.84, 0.82	ICC 0.94 (0.72–0.99)	*Moderate* *Good* *Moderate* *Good* *Good* *Moderate*

Table 2. *Cont.*

Device	Author	Intra-Rater Reliability	Inter-Rater Reliability	Level of Reliability
Adduction	Hawi et al. (2014) Hwang et al. (2017) Cai et al. (2019)	ICC 0.99 ICC 0.92 (0.89–0.93), 0.96 (0.86–0.96) ICC 0.70, 0.75, 0.84, 0.82		*Good* *Good* *Moderate*
External rotation	Huber et al. (2015) Çubukçu et al. (2020) Cai et al. (2019)	ICC 0.98 ICC 0.87 ICC 0.93, 0.75, 0.90, 0.60		*Good* *Good* *Moderate–Good*
Internal rotation	Çubukçu et al. (2020) Cai et al. (2019)	ICC 0.97 ICC 0.93, 0.75, 0.90, 0.60		*Good* *Moderate–Good*
Microsoft Kinect & Inertial Sensor				
Shoulder				
Flexion	Beshara et al. (2016)	ICC 0.84 (0.45–0.96), 0.93 (0.72–0.98)		*Moderate–Good*
Abduction	Beshara et al. (2016)	ICC 0.52 (-0.17–0.87, 0.85 (0.47–0.96)		*Poor–Moderate*
Inertial Sensor				
Shoulder				
Flexion	Rigoni et al. (2019) Chen et al. (2020) De Baets et al. (2020)	ICC 0.68, 0.87, 0.91	ICC 0.88 (0.80–0.92) ICC 0.90 (0.83–0.94), 0.95 (0.92–0.97) ICC 0.74, 0.83, 0.84	*Good* *Good* *Moderate–Good (Intra-rater)* *Moderate (Inter-rater)*
Extension	Chen et al. (2020) De Baets et al. (2020)	ICC 0.68, 0.87, 0.91	ICC 0.77 (0.64–0.87), 0.80 (0.68–0.89) ICC 0.74, 0.83, 0.84	*Moderate* *Moderate–Good (Intra-rater)* *Moderate (Inter-rater)*
Abduction	Cuesta-Vargas et al. (2016) Picerno et al. (2015) Rigoni et al. (2019) Chen et al. (2020) De Baets et al. (2020)	ICC 0.78 (0.40–0.93), 0.98 (0.94–0.99) 0.99 (0.98–0.99), 0.96 (0.93–0.98) ICC 0.96 (0.93–0.98) ICC 0.73, 0.95	ICC 0.49 (0.08–0.82), 0.99 (0.98–1.00), 0.99 (0.99–1.00) ICC 0.88 (0.81–0.93) ICC 0.97 (0.95–0.98), 0.98 (0.96–0.99) ICC 0.74, 0.80, 0.93	*Moderate–Good (Intra-rater)* *Poor–Good (Inter-rater)* *Good* *Good* *Good* *Moderate–Good*
Adduction	De Baets et al. (2020)	ICC 0.73, 0.95	ICC 0.74, 0.80, 0.93	*Moderate–Good*

Table 2. *Cont.*

Device	Author	Intra-Rater Reliability	Inter-Rater Reliability	Level of Reliability
External rotation	Schiefer et al. (2015) Rigoni et al. (2019) Chen et al. (2020) De Baets et al. (2020)	ICC 0.85, 0.87, 0.89, 0.90	ICC 0.71, 0.76, 0.81, 0.86 ICC 0.84 (0.74–0.90) ICC 0.95 (0.92–0.97), 0.96 (0.93–0.98) ICC 0.38, 0.84, 0.73, 0.87	*Moderate–Good* *Good* *Good* *Moderate–Good (Intra-rater)* *Poor–Good (Inter-rater)*
Internal rotation	Schiefer et al. (2015) Rigoni et al. (2019) Chen et al. (2020) De Baets et al. (2020)	ICC 0.85, 0.87, 0.89, 0.90	ICC 0.68, 0.78, 0.87, 0.98 ICC 0.71 (0.56–0.82) ICC 0.91 (0.86–0.95), 0.97 (0.94–0.98) ICC 0.38, 0.84, 0.73, 0.87	*Moderate–Good* *Moderate* *Good* *Moderate–Good (Intra-rater)* *Poor–Good (Inter-rater)*
Smartphone/Mobile App				
Shoulder				
Flexion	Chan et al. (2010) Shin et al. (2012) Werner et al. (2014) Mejia-Hernandez et al. (2018) Ramos et al. (2019)	ICC 0.99 ICC 0.97 (0.95–0.99), 0.96 (0.92–0.98) 0.99 (0.98–0.99), 0.99 (0.99–1.00) ICC −0.21, −0.19, 0.01, 0.16, 0.27, 0.40 0.47, 0.50, 0.53, 0.56, 0.60, 0.71, 0.76, 0.82	ICC 0.99 ICC 0.73 (0.59–0.83), 0.74 (0.61–0.84), 0.83 (0.73–0.90), 0.84 (0.74–0.90) ICC 0.75 (0.61–0.84), 0.97 (0.90–0.99) ICC 0.99 (0.98–0.99) ICC 0.06, 0.18, 0.19, 0.22, 0.25, 0.27, 0.30, 0.36, 0.40, 0.44, 0.47, 0.49, 0.68, 0.69	*Good* *Good (intra-rater)* *Moderate–Good (inter-rater)* *Moderate–Good* *Good* *Poor–Moderate*
Abduction	Lim et al. (2015) Shin et al. (2012) Werner et al. (2014) Mejia-Hernandez et al. (2018)	ICC 0.72, 0.89, 0.95, 0.97 ICC 0.96, 0.97, 0.99	ICC 0.79, 0.94 ICC 0.70, 0.72, 0.78, 0.79 ICC 0.72 (0.58–0.83), 0.91 (0.68–0.97) ICC 0.99 (0.99–1.00)	*Moderate–Good* *Good (intra-rater)* *Moderate (inter-rater)* *Moderate–Good* *Good*
Glenohumeral abduction	Mejia-Hernandez et al. (2018)		ICC 0.98 (0.97–0.99), 0.97 (0.95–0.99)	*Good*

Table 2. *Cont.*

Device	Author	Intra-Rater Reliability	Inter-Rater Reliability	Level of Reliability
External rotation	Chan et al. (2010) Mitchell et al. (2014) Shin et al. (2012) Werner et al. (2014) Mejia-Hernandez et al. (2018)	ICC 0.94, 0.96 ICC 0.79 (0.70–0.86) ICC 0.95, 0.97, 0.98	ICC 0.88, 0.98 ICC 0.94 (0.87–0.98) ICC 0.76, 0.77, 0.78, 0.89, 0.90 ICC 0.85 (0.57–0.95), 0.86 (0.79–0.92), 0.88 (0.66–0.96) ICC 0.99 (0.98–0.99)	*Good* *Moderate (intra-rater)* *Good (inter-rater)* *Good (intra-rater)* *Moderate–Good (inter-rater)* *Good* *Good*
Internal rotation	Shin et al. (2012) Werner et al. (2014) Mejia-Hernandez et al. (2018)	ICC 0.79, 0.97, 0.90, 0.93 0.99	ICC 0.63, 0.66, 0.67, 0.68 ICC 0.81 (0.70–0.88), 0.86 (0.48–0.93) ICC 0.98 (0.97–0.99), 0.98 (0.96–0.98)	*Moderate–Good* *Good* *Good*
Scaption	Ramos et al. (2019)	ICC −0.04, 0.10, 0.12, 0.31, 0.32, 0.39, 0.40, 0.45, 0.47, 0.52, 0.57, 0.63	ICC −0.17, −0.06, 0.03, 0.07, 0.23, 0.26, 0.27, 0.28, 0.36, 0.45, 0.54, 0.73, 0.75, 0.81	*Poor–Moderate (intra-rater)* *Poor–Good (inter-rater)*
Digital Inclinometer/Goniometer				
Shoulder				
Flexion	Dougherty et al. (2015) Kolber et al. (2011) Scibek et al. (2013) Correll et al. (2018)	ICC 0.77, 0.82 ICC 0.83 ICC 0.67, 0.80, 0.90, 0.92, 0.96 ICC 0.86, 0.88	ICC 0.58 ICC 0.18, 0.33, 0.50, 0.62, 0.68, 0.72, 0.76, 0.78, 0.85 ICC 0.89	*Moderate* *Moderate (intra-rater)* *Poor (inter-rater)* *Moderate–Good (intra-rater)* *Poor–Moderate (inter-rater)* *Good*
Elevation	Walker et al. (2016)	ICC 0.91, 0.92, 0.93, 0.95		*Good*
Glenohumeral flexion	Dougherty et al. (2015) Scibek et al. (2013)	ICC 0.75, 0.77 ICC 0.75, 0.92, 0.94, 0.96, 0.99	ICC 0.14, 0.35, 0.43, 0.63, 0.65, 0.69, 0.72, 0.83	*Moderate* *Moderate–Good (intra-rater)* *Poor–Good (inter-rater)*

Table 2. *Cont.*

Device	Author	Intra-Rater Reliability	Inter-Rater Reliability	Level of Reliability
Abduction	deWinter et al. (2004) Kolber et al. (2011) Poser et al. (2015) Dougherty et al. (2015) Scibek et al. (2013) Walker et al. (2016) Correll et al. (2018)	ICC 0.91 ICC 0.83, 0.92, 0.93, 0.96 ICC 0.73, 0.76 ICC 0.91, 0.94, 0.95, 0.96, 0.97, 0.99 ICC 0.89, 0.90, 0.91, 0.94 ICC 0.86, 0.91	ICC 0.28, 0.78, 0.83 ICC 0.95 ICC 0.27, 0.32, 0.40, 0.60, 0.63, 0.64 ICC 0.48, 0.56, 0.58, 0.62, 0.65, 0.68, 0.70, 0.80, 0.83 ICC 0.93	*Poor–Good* *Moderate* *Moderate–Good (intra-rater)* *Poor–Moderate (inter-rater)* *Moderate* *Good (intra-rater)* *Poor–Good (inter-rater)* *Good* *Good*
Glenohumeral abduction	Dougherty et al. (2015)	ICC 0.60, 0.75		*Moderate*
Horizontal abduction	Poser et al. (2015)	ICC 0.66, 0.81, 0.91, 0.94, 0.96	ICC 0.17, 0.18, 0.24, 0.28, 0.31	*Moderate–Good (intra-rater)* *Poor (inter-rater)*
Digital Inclinometer/Goniometer External rotation	Awan et al. (2002) Cools et al. (2014) deWinter et al. (2004) Kolber et al. (2011) Poser et al. (2015) Dougherty et al. (2015) Walker et al. (2016) Correll et al. (2018)	ICC 0.58, 0.67 ICC 0.98, 0.95, 0.98 ICC 0.94 ICC 0.93, 0.94, 0.96, 0.97 ICC 0.28, 0.61, 0.66, 0.64, 0.68, 0.71 ICC 0.90, 0.94, 0.95 ICC 0.89, 0.90	ICC 0.41, 0.51 ICC 0.98 ICC 0.56, 0.88, 0.90 ICC 0.88 ICC 0.70, 0.71, 0.72, 0.73, 0.76, 0.77 ICC 0.98	*Poor–Moderate (intra-rater)* *Poor (inter-rater)* *Good* *Poor–Good* *Good* *Good (intra-rater)* *Moderate (inter-rater)* *Poor–Moderate* *Good* *Good*
Internal rotation	Awan et al. (2002) Cools et al. (2014) Kolber et al. (2011) Poser et al. (2015) Dougherty et al. (2015) Walker et al. (2016) Correll et al. (2018)	ICC 0.64, 0.65, 0.72 ICC 0.89, 0.98, 0.99 ICC 0.87 ICC 0.91, 0.92, 0.94, 0.96, 0.97 ICC 0.64, 0.68 ICC 0.85, 0.90, 0.93, 0.96 ICC 0.82, 0.85	ICC 0.50, 0.52, 0.62, 0.66 ICC 0.96, 0.98 ICC 0.93 ICC 0.63, 0.66, 0.73, 0.76, 0.78 ICC 0.96	*Poor–Moderate* *Good* *Good* *Good (intra-rater)* *Moderate (inter-rater) Poor–Moderate* *Moderate–Good* *Moderate–Good*
Scaption	Kolber et al. (2012)	ICC 0.88	ICC 0.89	*Good*

ICC = intra-class correlation coefficient. Level of reliability determined by the criteria identified by Swinkels et al. [87].

3.3.1. The Microsoft Kinect

Six studies assessed intra-rater reliability [93,94,99,105–107], one study assessed inter-rater reliability [112] and another study assessed both [101]. Two studies reported good intra-rater reliability (ICC > 0.85) for all shoulder movements [105,107]. The remaining four studies reported varying levels of intra-rater reliability, ranging from poor (ICC < 0.65), moderate (ICC 0.65–0.85) to good, dependent on the shoulder movements assessed. Shoulder external and internal rotation demonstrated moderate to good levels of intra-rater reliability [94,99,106]. Two studies reported good inter-rater reliability (ICC > 0.80) for shoulder flexion, extension, and abduction [101,112]. Intra-rater reliability for coupling inertial sensors with the Kinect was moderate to good for flexion, and poor to moderate for abduction [92].

3.3.2. Inertial Sensors

One study assessed intra-rater reliability [114], three studies assessed inter-rater reliability [96,117,118] and two studies assessed both [100,102]. Three studies reported moderate to good intra-rater reliability using one, two or four wearable inertial sensors [100,102,114]. Inter-rater reliability was good or moderate in four studies for shoulder flexion, extension, abduction, external and internal rotation [96,102,117,118]. A wider range of poor (ICC < 0.60) to good inter-rater reliability was reported in two studies for shoulder abduction, external and internal rotation [100,102].

3.3.3. Smartphone/Mobile Applications

A total of five of seven studies [95,110,113,116,120] assessed intra-rater and inter-rater reliability. All shoulder movements across most of the studies demonstrated moderate or good levels of intra- and inter-rater reliability. Only one study reported a wider range of reliability values, between poor and good, for flexion and scaption [116].

3.3.4. Digital Inclinometer/Goniometer

Two studies assessed intra-rater reliability [104,121], one study assessed inter-rater reliability [103], and seven studies assessed both [91,97,98,108,109,115,119]. Intra-rater reliability was predominately moderate to good for all shoulder movements (n = 7). Two studies reported poor to moderate intra-rater reliability for external and internal rotation [91,104]. Poor inter-rater reliability was reported in five studies [91,103,108,115,119]. Only two studies reported good intra- and inter-rater reliability for all shoulder movements [97,109].

3.4. Methodological Evaluation of the Measurement Properties

Of the thirty-two included studies, only two [109,110] scored very good on all items of the COSMIN reliability standards checklist. A total of 24 studies scored adequate, five were rated doubtful and one was rated inadequate. Table 3 lists the COSMIN standards of reliability checklist and all subsequent scores.

Using the COSMIN criteria, only one study [109] was found to have a very good score on all items for the measurement error standards. A total of 18 studies scored adequate, with two rated doubtful and 11 rated inadequate. Table 4 lists all items of the COSMIN standards on measurement error checklist and the subsequent paper scores.

Table 3. Assessment of reliability using the COSMIN standards for studies on reliability checklist.

Items	First Author and Year										
	Awan (2002)	**Beshara (2016)**	**Bonnechère (2014)**	**Cai (2019)**	**Chan (2010)**	**Chen (2020)**	**Cools (2014)**	**Correll (2018)**	**Çubukçu (2020)**	**Cuesta-Vargas (2016)**	**Da Cunha Neto (2018)**
1. Were patients stable in the time between repeated measurements on the construct to be measured?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG	A
2. Was the time interval between the repeated measurements appropriate?	A	VG	VG	VG	A	A	A	A	VG	VG	A
3. Were the measurement conditions similar for the repeated measurements–except for the condition being evaluated as a source of variation?	VG	VG	VG	VG	A	VG	VG	VG	VG	VG	A
4. Did the professional(s) administer the measurement without knowledge of scores or values of other repeated measurement(s) in the same patients?	VG	VG	VG	A	VG	VG	VG	VG	A	A	A
5. Did the professionals(s) assign scores or determine values without knowledge of scores or values of other repeated measurements(s) in the same patients?	VG	VG	VG	A	VG	VG	VG	VG	A	A	A
6. Were there any other important flaws in the design or statistical methods of the study?	D	D	VG	A	I	A	VG	VG	VG	VG	A
7. For continuous scores: was an intraclass correlation (ICC) calculated?	A	VG	A	VG	A	VG	VG	VG	VG	VG	A
8. For ordinal scores: was a (weighted) kappa calculated?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9. For dichotomous/nominal scores: was Kappa calculated for each category against the other categories combined?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Overall Score	D	D	A	A	I	A	A	A	A	A	A

Table 3. *Cont.*

Items	First Author and Year										
	De Baets (2020)	**deWinter (2004)**	**Dougherty (2015)**	**Hawi (2014)**	**Huber (2015)**	**Hwang (2017)**	**Kolber (2011)**	**Kolber (2012)**	**Lim (2015)**	**Mejia-Hernandez (2018)**	**Milgrom (2016)**
1 Were patients stable in the time between repeated measurements on the construct to be measured?	VG	VG	A	VG	VG	VG	VG	VG	VG	VG	VG
2. Was the time interval between the repeated measurements appropriate?	A	A	VG	A	A	A	A	VG	VG	A	A
3. Were the measurement conditions similar for the repeated measurements–except for the condition being evaluated as a source of variation?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG
4. Did the professional(s) administer the measurement without knowledge of scores or values of other repeated measurement(s) in the same patients?	A	A	VG	A	VG	A	VG	VG	VG	VG	A
5. Did the professionals(s) assign scores or determine values without knowledge of scores or values of other repeated measurements(s) in the same patients?	A	A	VG	A	VG	A	VG	VG	VG	VG	A
6. Were there any other important flaws in the design or statistical methods of the study?	A	VG	VG	D	A	D	VG	VG	VG	VG	D
7. For continuous scores: was an intraclass correlation (ICC) calculated?	VG	A	A	VG	VG	A	VG	VG	VG	A	A
8. For ordinal scores: was a (weighted) kappa calculated?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9. For dichotomous/nominal scores: was Kappa calculated for each category against the other categories combined?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Overall Score	A	A	A	D	A	D	A	VG	VG	A	D

Table 3. *Cont.*

Items	First Author and Year									
	Mitchell (2014)	**Picerno (2015)**	**Poser (2015)**	**Ramos (2019)**	**Rigoni (2019)**	**Schiefer (2015)**	**Scibek (2013)**	**Shin (2012)**	**Walker (2016)**	**Werner (2014)**
1. Were patients stable in the time between repeated measurements on the construct to be measured?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG
2. Was the time interval between the repeated measurements appropriate?	A	A	VG	VG	A	A	A	A	A	A
3. Were the measurement conditions similar for the repeated measurements–except for the condition being evaluated as a source of variation?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG
4. Did the professional(s) administer the measurement without knowledge of scores or values of other repeated measurement(s) in the same patients?	VG	A	A	A	VG	VG	A	VG	VG	VG
5. Did the professionals(s) assign scores or determine values without knowledge of scores or values of other repeated measurements(s) in the same patients?	VG	A	A	A	VG	VG	A	VG	VG	VG
6. Were there any other important flaws in the design or statistical methods of the study?	VG	VG	A	VG	VG	A	A	VG	A	A
7. For continuous scores: was an intraclass correlation (ICC) calculated?	VG	VG	VG	A	VG	VG	VG	VG	VG	VG
8. For ordinal scores: was a (weighted) kappa calculated?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9. For dichotomous/nominal scores: was Kappa calculated for each category against the other categories combined?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Overall Score	A	A	A	A	A	A	A	A	A	A

Abbreviations: VG: very good; A: adequate, D: doubtful; I: inadequate; N/A: not applicable.

Table 4. Assessment of measurement error using the COSMIN standards for studies on measurement error checklist.

Items	First Author and Year										
	Awan (2002)	**Beshara (2016)**	**Bonnechère (2014)**	**Cai (2019)**	**Chan (2010)**	**Chen (2020)**	**Cools (2014)**	**Correll (2018)**	**Çubukçu (2020)**	**Cuesta-Vargas (2016)**	**Da Cunha Neto (2018)**
1 Were patients stable in the time between repeated measurements on the construct to be measured?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG	A
2. Was the time interval between the repeated measurements appropriate?	A	VG	A	VG	A	A	A	A	VG	VG	D
3. Were the measurement conditions similar for the repeated measurements–except for the condition being evaluated as a source of variation?	VG	VG	VG	VG	A	VG	VG	VG	VG	VG	A
4. Did the professional(s) administer the measurement without knowledge of scores or values of other repeated measurement(s) in the same patients?	VG	VG	VG	A	VG	VG	VG	VG	A	A	A
5. Did the professionals(s) assign scores or determine values without knowledge of scores or values of other repeated measurements(s) in the same patients?	VG	VG	VG	A	VG	VG	VG	VG	A	A	A
6 Were there any other important flaws in the design or statistical methods of the study?	D	D	VG	D	I	VG	VG	VG	VG	VG	VG
7. For continuous scores: was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC), Limits of Agreement (LoA) or Coefficient of Variation (CV) calculated?	I	VG	VG	I	I	VG	VG	VG	VG	I	I
8. For dichotomous/nominal/ordinal scores: was the percentage specific (e.g., positive and negative) agreement calculated?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Overall Score	I	D	A	I	I	A	A	A	A	I	I

Table 4. *Cont.*

Items	First Author and Year										
	De Baets (2020)	**deWinter (2004)**	**Dougherty (2015)**	**Hawi (2014)**	**Huber (2015)**	**Hwang (2017)**	**Kolber (2011)**	**Kolber (2012)**	**Lim (2015)**	**Mejia-Hernandez (2018)**	**Milgrom (2016)**
1. Were patients stable in the time between repeated measurements on the construct to be measured?	VG	VG	A	VG	VG	VG	VG	VG	VG	VG	VG
2. Was the time interval between the repeated measurements appropriate?	A	A	VG	A	A	A	A	VG	VG	A	A
3. Were the measurement conditions similar for the repeated measurements–except for the condition being evaluated as a source of variation?	VG	VG	VG	A	VG	VG	VG	VG	VG	VG	VG
4. Did the professional(s) administer the measurement without knowledge of scores or values of other repeated measurement(s) in the same patients?	A	A	VG	A	VG	A	VG	VG	VG	VG	A
5. Did the professionals(s) assign scores or determine values without knowledge of scores or values of other repeated measurements(s) in the same patients?	A	A	VG	A	VG	A	VG	VG	VG	VG	A
6. Were there any other important flaws in the design or statistical methods of the study?	VG	VG	VG	D	VG	D	VG	VG	D	VG	D
7. For continuous scores: was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC), Limits of Agreement (LoA) or Coefficient of Variation (CV) calculated?	VG	N/A	VG	I	VG	VG	VG	VG	I	VG	I
8. For dichotomous/nominal/ordinal scores: was the percentage specific (e.g., positive and negative) agreement calculated?	N/A	VG	A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Overall Score	A	A	A	I	A	D	A	VG	I	A	I

Table 4. *Cont.*

Items	First Author and Year									
	Mitchell (2014)	**Picerno (2015)**	**Poser (2015)**	**Ramos (2019)**	**Rigoni (2019)**	**Schiefer (2015)**	**Scibek (2013)**	**Shin (2012)**	**Walker (2016)**	**Werner (2014)**
1. Were patients stable in the time between repeated measurements on the construct to be measured?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG
2. Was the time interval between the repeated measurements appropriate?	A	A	VG	VG	A	A	A	A	A	A
3. Were the measurement conditions similar for the repeated measurements–except for the condition being evaluated as a source of variation?	VG	VG	VG	VG	VG	VG	VG	VG	VG	VG
4. Did the professional(s) administer the measurement without knowledge of scores or values of other repeated measurement(s) in the same patients?	VG	A	A	A	VG	VG	A	VG	VG	VG
5. Did the professionals(s) assign scores or determine values without knowledge of scores or values of other repeated measurements(s) in the same patients?	VG	A	A	A	VG	VG	A	VG	VG	VG
6. Were there any other important flaws in the design or statistical methods of the study?	VG	VG	VG	VG	VG	VG	D	VG	VG	VG
7. For continuous scores: was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC), Limits of Agreement (LoA) or Coefficient of Variation (CV) calculated?	I	I	VG	VG	VG	VG	I	VG	VG	VG
8. For dichotomous/nominal/ordinal scores: was the percentage specific (e.g., positive and negative) agreement calculated?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Overall Score	I	I	A	A	A	A	I	A	A	A

Abbreviations: VG: very good; A: adequate, D: doubtful; I = inadequate; N/A: not applicable.

3.5. Synthesis of Results (Meta-Analysis)

ICC values were included from studies with n > 1 participant included in intra- and inter-rater reliability analysis. The ICC values for outcome measures (aROM or pROM for abduction, flexion, internal rotation, external rotation) were individually assessed based on motion and grouped by method (K, SP, DG, DI and IS) to produce a pooled correlation with a 95%CI (Figures 2–4).

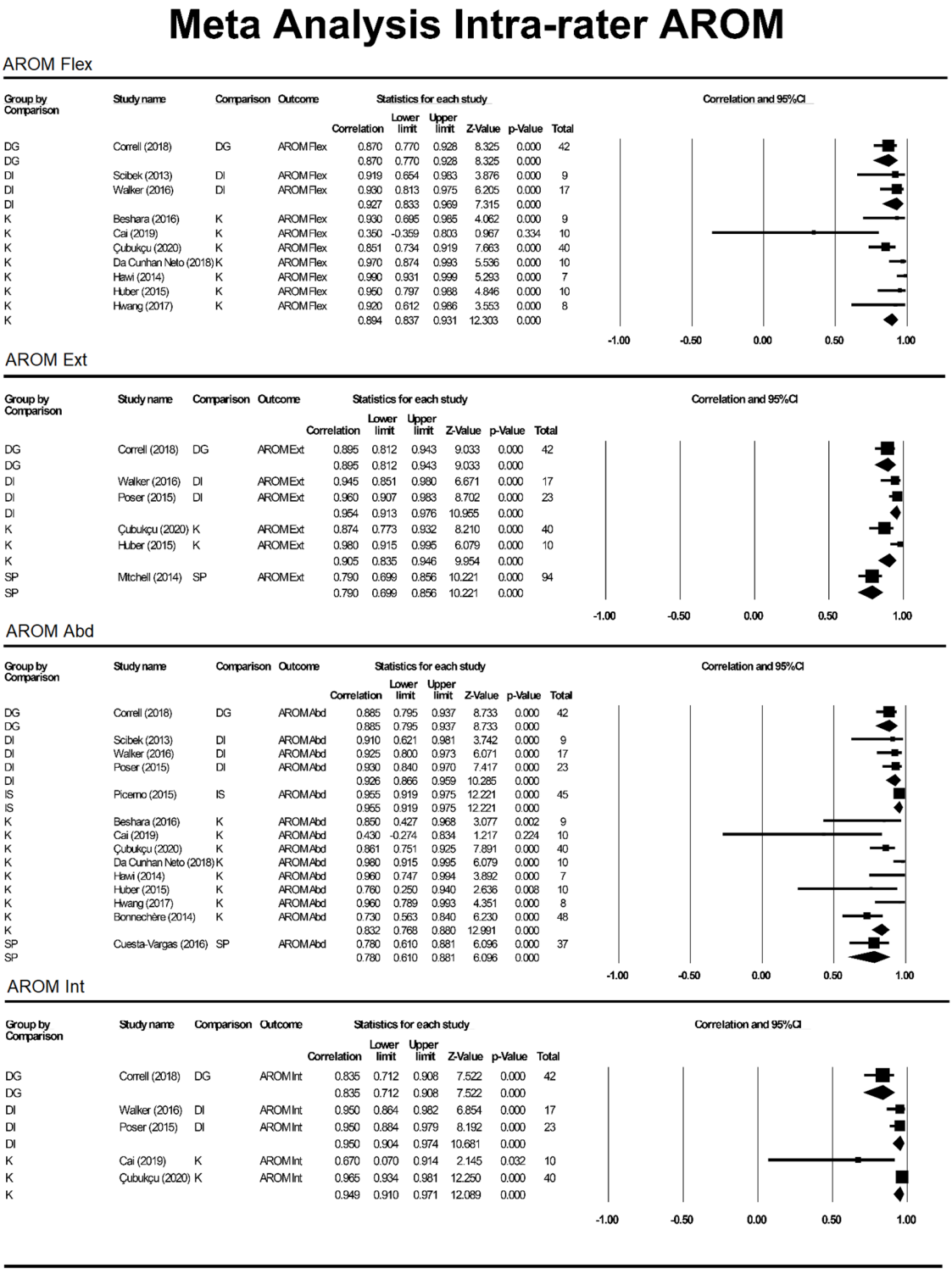

Figure 2. Meta-analysis for intra-rater AROM. DG = digital goniometer, DI = digital inclinometer, IS = inertial sensor, K = Kinect, SP = smartphone, Abd = abduction, AROM = active range of motion, Int = internal rotation, Ext = external rotation.

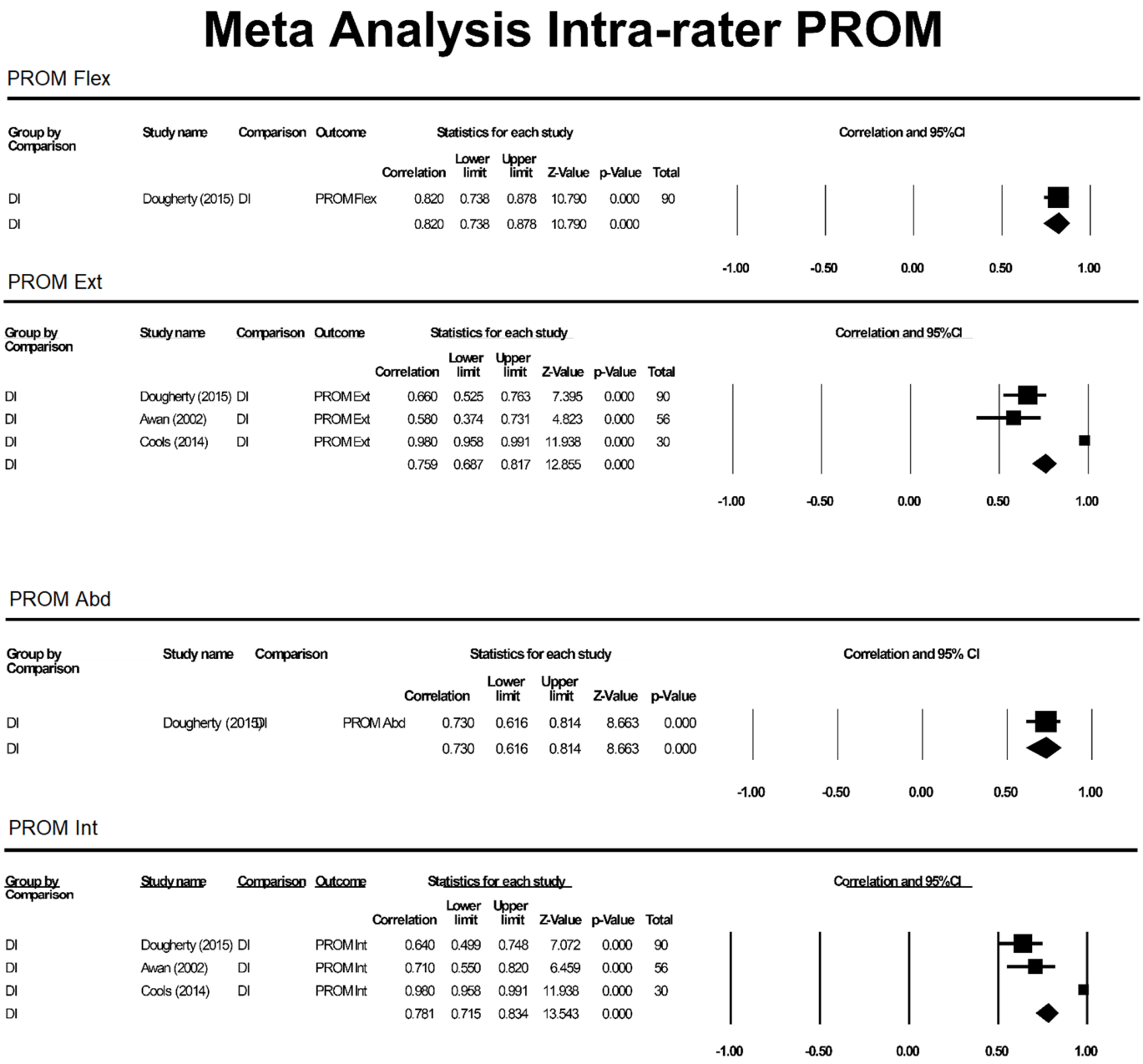

Figure 3. Meta-analysis for intra-rater pROM. DI = digital inclinometer, PROM = passive range of motion, Flex = flexion, Int = internal rotation, Ext = external rotation.

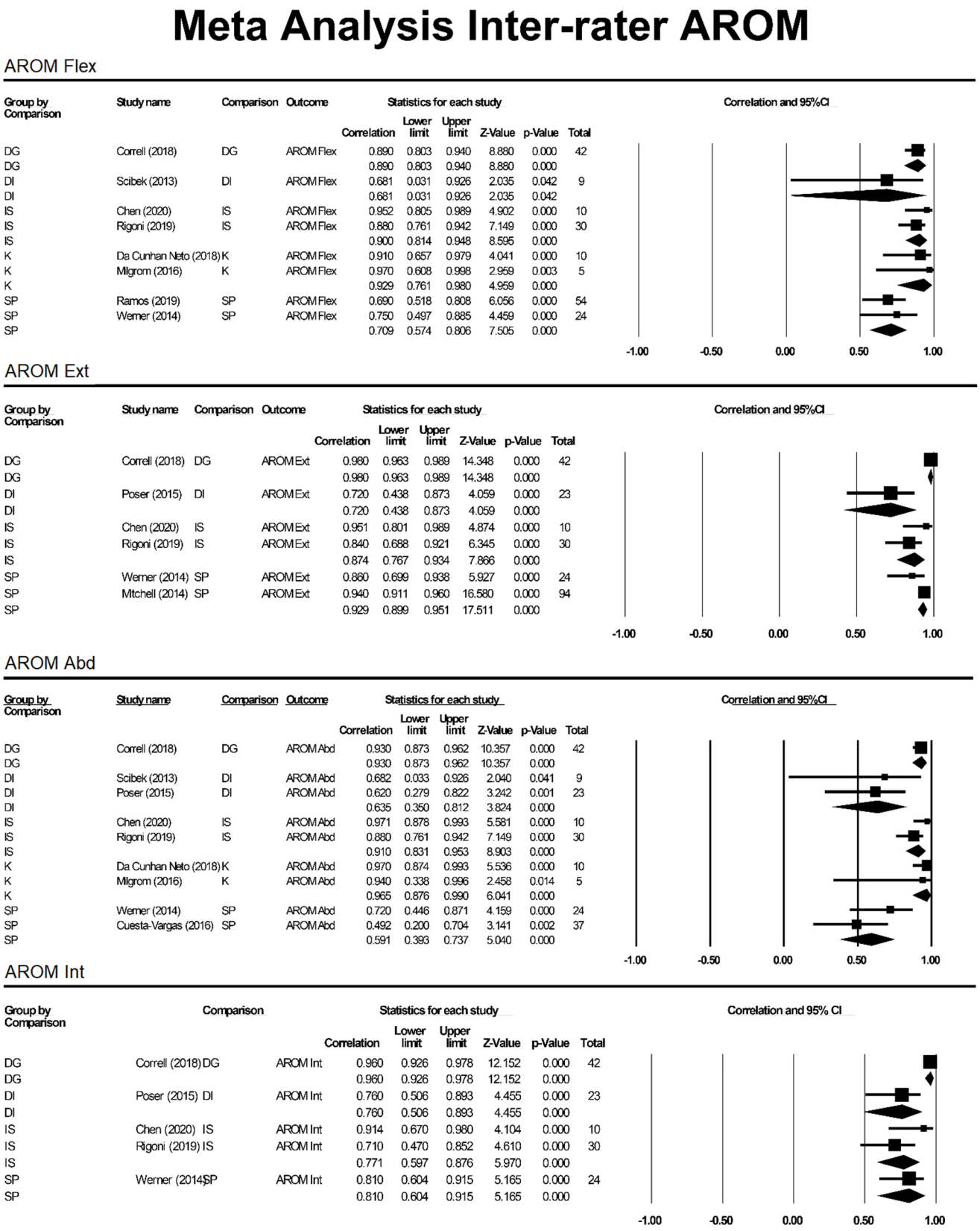

Figure 4. Meta-analysis for inter-rater aROM. DG = digital goniometer, DI = digital inclinometer, IS = inertial sensor, K = Kinect, SP = smartphone, Abd = abduction, Int = internal rotation.

3.6. Anatomical Landmarks

Twenty-three studies identified the anatomical landmarks for each device and are summarised in Table 5. A total of six studies reported using a vector from the shoulder joint to the elbow for the Microsoft Kinect [92–94,105].

Five studies identified the anatomical landmarks for inertial sensor placement [100,102,114,117,118]. All studies used a sensor located on the upper arm that was either unspecified (n = 2), placed on the middle third of the humerus (n = 3), or attached 10 cm distal to the lateral epicondyle (n = 1). Two studies placed a sensor on the flat part of the sternum [100,102]. Only two studies reported using a lower arm sensor on the wrist [102,118].

Anatomical positions for smartphone device placement were described in five studies [95,110,111,113,120]. The most common attachment was on the humerus (n = 3) followed by positions at the wrist (n = 2).

Seven studies reported anatomical landmarks for digital inclinometers [98,104,108,109,115,119,121]. Locations were predominately determined by the type of shoulder movement performed, orientation, and assessment position.

Table 5. Anatomical landmarks by device.

Device	Author	Anatomical Landmarks
Microsoft Kinect	Bonnechère et al. (2014)	Shoulder girdle centre, elbow, wrist, hand
	Hawi et al. (2014)	Shoulder centre and elbow
	Huber et al. (2015)	Positions of shoulder and elbow joints relative to the trunk for flexion and abduction. Position of elbow and hand relative to trunk for external rotation
	Milgrom et al. (2016)	Angle between the humerus vector (shoulder to elbow) and the torso vector (neck to shoulder midpoint)
	Cai et al. (2019)	X = Unit vector perpendicular to the Y-axis and the Z-axis pointing anteriorly, Y: Unit vector going from the elbow joint center to the shoulder joint center, Z: Unit vector perpendicular to the plane formed by the Y-axis of the upper arm and the long axis vector of the forearm.
Microsoft Kinect & Inertial Sensor	Beshara et al. (2016)	2 3D vectors, a vector from shoulder joint centre (below the acromion process) to the elbow centre (between the medical and lateral epicondyles). A vector from shoulder joint centre defined as a point on the 6th rib along the midaxillary line of the trunk.
Inertial Sensor	Cuesta-Vargas et al. (2016)	Middle third of the humerus slightly posterior and in the flat part of the sternum
	Picerno et al. (2015)	Arbitrary point of the upper arm
	Schiefer et al. (2015)	Laterally on the upper arms and on the forearms close to the wrist, on the dorsum of the hand. Sensors were placed in the middle of the segments.
	Rigoni et al. (2019)	10 cm distal to the lateral epicondyle
	De Baets et al. (2020)	*Sternal sensor:* positioned on flat central part of the sternum, the *scapular sensor* halfway between the trigonum and the acromial angle, in alignment with the upper edge of the scapular spine. *Humeral sensor:* at the central third of the humerus, slightly posterior, at the level of the deltoid insertion. *Lower arm sensor:* positioned on the dorsal side, just proximal of the line between the radial and ulnar styloid process.

Table 5. *Cont.*

Device	Author	Anatomical Landmarks
Smartphone/mobile app	Chan et al. (2010)	Acromion, humeral axis.
	Lim et al. (2015)	Front centre of humerus.
	Mitchell et al. (2014)	Superior border of the mid-ulna.
	Ramos et al. (2019)	Attached below the deltoid muscle origin
	Mejia-Hernandez et al. (2018)	Distal portion of the humerus for seated movements. Wrist for supine movements.
	Shin et al. (2012)	Ventral side of the patient's forearm at the wrist level.
Digital Inclinometer	Dougherty et al. (2015)	*Shoulder flexion:* the anterior aspect of the arm, aligned parallel to the humerus. *Shoulder abduction:* The lateral aspect of the arm, aligned parallel to the humerus. *External Rotation:* The anterior aspect of the participant's forearm. *Internal Rotation:* The posterior aspect of the participant's forearm.
	Kolber et al. (2011)	*Flexion:* Distal arm proximal to the elbow. *Abduction:* Distal arm proximal to the elbow. *External rotation:* Distal forearm just proximal to the wrist. *Internal rotation:* Distal forearm just proximal to the wrist.
	Kolber et al. (2012)	*Scaption:* Superior portion of the humeral shaft proximal to the elbow.
	Poser et al. (2015)	*Abduction:* lateral and distal face of the humerus, with the inferior edge set at the beginning of the medial epicondyle. *Horizontal adduction:* spine of scapular and posterior face of the humerus, touching the olecranon.
	Scibek et al. (2013)	*Flexion and Abduction:* Shaft of the humerus
	Walker et al. (2016)	*Shoulder internal/external rotation:* 5 cm distal to the olecranon process of the elbow. *Combined elevation:* Just below the deltoid insertion with the face of the inclinometer in the coronal plane of movement. *Shoulder abduction in internal rotation:* Just below the deltoid insertion with the face of the inclinometer in the coronal plane of movement.
	Correll (2018)	*Shoulder flexion:* the lateral aspect of the greater tubercle, the midaxillary line of the thorax and the lateral midline of the humerus, lateral epicondyle of the humerus or the olecranon process. *Abduction:* anterior aspect of the acromial process, midline of the anterior aspect of the sternum and the anterior midline of the humerus. *External/Internal rotation:* the olecranon process, the ulna and ulnar styloid.

4. Discussion

Thirty-two studies investigating four different types of devices were included in this review. A thorough search of relevant literature found no previous systematic review of intra-rater and inter-rater reliability of the Microsoft Kinect and ambulatory sensor-based motion tracking devices to measure shoulder ROM.

Good intra-rater reliability for multiple types of shoulder movement was demonstrated with the Kinect [105,107], smartphone applications [95], and digital inclinometers [97]. The Kinect consistently demonstrated higher intra-rater ICC values over other

devices for all shoulder movements. Only one study reported poor intra-rater reliability for measuring shoulder extension with the Kinect [99]. Overall, inertial sensors, smartphones, and digital inclinometers demonstrated moderate to good intra-rater reliability across all shoulder movements.

Good inter-rater reliability for more than one type of shoulder movement was demonstrated with the Kinect [101,112], smartphone applications [95,111], and a digital inclinometer/goniometer [97,98]. Inertial sensors predominately exhibited moderate to good inter-rater reliability across all types of shoulder movements. Broader ranges of inter-rater reliability (between poor to moderate) were more commonly reported with digital goniometers.

4.1. Quality of Evidence

All included studies and measurement properties were assessed for their methodological quality using the COSMIN tool. The methodological quality ranged from doubtful to very good for reliability standards. The strict COSMIN criteria of using the worst-score counts to denote the overall score resulted in only two very good studies [109,110], which reported moderate or good reliability for using a digital inclinometer and a smartphone device. An adequate rating was scored by five studies for the Kinect, six studies for inertial sensors, five studies for smartphone applications, and eight studies for digital inclinometers/goniometers.

Five studies missed achieving an overall very good rating due to receiving only an adequate score for the time interval between measurements (COSMIN item two). The authors acknowledge an appropriate time interval depends on the stability of the construct (COSMIN item one), and the target population [88]. The time interval must be adequately distanced to avoid recall bias, yet within a compact enough window to distinguish genuine differences in measurements from clinical change [123–125]. Studies had a time interval ranging from the same day (22/32) to 7 days (4/32) between two repeated measurements. Ideal time intervals of 2–7 days have been recommended to minimise the risk of a learning effect, random error, or other modifying factors that can affect the movement pattern [126,127].

Small sample sizes contributed to five studies scoring doubtful or inadequate, in accordance with COSMIN item six. An insufficient sample size may not detect true differences and reduces the power of the study to draw conclusions [128]. Of the 32 included studies, a power analysis for sample size calculation was reported in only four (12.5%) studies [97,98,100,104]. The latest COSMIN checklist has removed the standards for adequate sample sizes, as the authors suggested that several small high-quality studies can together provide good evidence for the measurement property [129]. The guidelines recommend a more nuanced approach that considers several factors including the type of ICC model. Studies with small sample sizes were considered acceptable if the authors justified the reasons outlining its adequacy [129]. Therefore, for methodological quality, reviewers scored sample sizes of 1 inadequate, <10 doubtful, <30 adequate and $\geq$30 very good. This criterion was based on literature citing a rule of thumb of recruiting 19–30 participants when conducting a reliability study [130–132].

With respect to measurement error assessment, just over one-half of the studies (18/32) scored adequate, and one scored very good for methodological quality. Eleven studies were rated inadequate, as they failed to calculate SEM, SDC, LoA or CV values (COSMIN item seven). Two studies [92,107] were rated doubtful due to minor methodological flaws (COSMIN item six); notably, this strict item offered reviewers no adequate option.

Reliability and measurement error are inextricably linked, and a highly reliable measurement contains little measurement error. A clinician can confidently verify real changes in patient status if the measured change from the last measurement is larger than the error associated with the measurement [133]. The minimal detectable clinical difference at a 90% confidence level (MDC90) is the minimal value to determine whether a change has occurred [72]. MDC values are open to interpretation and are based on clinical judgement.

For shoulder ROM measurement, differences between observers which exceed 10° are deemed unacceptable for clinical purposes [103].

The Kinect and inertial sensors demonstrated low SEM and MDC values for measuring most types of shoulder movements [96,99,102,106]. Similarly, for the Kinect, low CV values (1.6%, 5.9%) were reported for shoulder abduction [93]. Smartphones had moderate SEM and MDC values, with better (smaller) errors demonstrated for intra-rater analysis [118], abduction and forward flexion [122], and higher target angles [116]. One study comparing smartphone measurements with universal goniometry, analysed Bland–Altman plots to indicate narrow LoA and excellent agreement, particularly for glenohumeral abduction [111]. Digital inclinometers demonstrated low MDC90 values, ranging from ≥2.82° to 5.47° [97], 4° to 9° [108], and 5° to 12° [121] for inter-rater analysis. Four studies reported acceptable differences between observers of <±10° [97,103,108,121] for most inclinometer measurements.

4.2. Clinical Implications

The Microsoft Kinect is an affordable depth imaging technology that can conveniently and reliably measure shoulder aROM. As a low-cost markerless system, the Kinect can provide clinicians with fast, real-time objective data to quantify shoulder kinematics. The Kinect's visual feedback can aid in patient motivation by way of monitoring treatment and disease progression. The massive amounts of kinematic data generated allows clinicians to potentially analyse shoulder motion paths and correlate specific movement patterns to shoulder pathology [94]. Moreover, higher clinical efficiency arises from relying less on time and labour-intensive patient-reported outcome measures. The portability of the Microsoft Kinect over expensive motion capture systems permits its practical use in private clinics, rehabilitation centres, and home settings [107].

All studies were limited to motion performed along the anatomical planes. The simplicity of calculating the angles between two corresponding vectors does not take into account movements that occur outside the plane. In contrast to goniometric measurements, Lee et al. [60] found subjects could abduct their shoulders to a greater degree in front of the Kinect because their movements were not controlled in a given anatomical plane by an examiner. The authors performed a supplementary experiment that compared goniometric and Kinect shoulder measurements in rapid succession within three cardinal planes. Results demonstrated a significant decrease in 95% limit of agreement between both methods in all directions. It was concluded that the variability was due to the unrestricted motion of the Kinect.

With respect to reliability, one study reported lower repeatability with the Kinect in the frontal and transverse planes compared to the sagittal plane [94]. Another study reported large discrepancies for precise shoulder angle measurements with the Kinect [106].

Discrepancies between standing, sitting, and lying positions can also be a source of difference for shoulder ROM measurements [134,135]. One study [60] reported discrepancies between goniometric shoulder ROM measurements with seated subjects and Kinect ROM measurements for standing subjects. The authors attributed this result to the limitation of the Kinect's skeletal tracking, which is optimised for standing rather than sitting. Moreover, better accuracy for the Kinect has been reported for standing postures [136]. Therefore, adequate patient positioning and protocol standardisations are essential to reduce measurement error [105]. Suggested examples include placing coloured footprints on the floor and fixating the Kinect sensor bar [105].

Wearability and usability are two aspects to consider for implementing sensor-based human motion tracking devices in clinical practice. For the included studies, IMUs were most often positioned on the upper arm with additional placements on the sternum and wrist. Methods of fixation included double-sided adhesive tape with an elastic cohesive [100], an elastic belt [114] and, velcro straps [118]. Smartphone devices were attached by commercial armbands [110,111,116,120] or were hand-held by the examin-

ers [95,113,122]. Most notably, no studies reported any calibration issues, and only one study [111] reported attachment difficulties.

4.3. Limitations

There were some limitations in this review. First, no additional search was performed for grey literature, and only studies written in English were included. Although authors identified an additional six reliability studies, they were excluded because they did not assess ICCs.

Second, the authors acknowledge the limitations of the revised COSMIN methodological quality tool, as it was primarily developed to assess risk of bias and not study design. Although more user-friendly than the original version, the omission of a sample size criterion leaves open a wider interpretation as to what constitutes an adequate sample size. Furthermore, no standards exist regarding the types of patients, examiners (well-trained or otherwise), and testing procedures. Future studies can apply other tools such as the modified GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach to address these issues [137]. Additionally, because the revised COSMIN guidelines are relatively new update, caution should be exercised when interpreting and comparing these results with prior studies that used the original COSMIN checklist.

Third, our meta-analysis was limited by the heterogeneity of the studies, given the variance in sample sizes, protocols, shoulder positions, and number of raters. Several studies did not report the 95% confidence intervals for ICCs. Furthermore, the calculation methods for ROM angles with the Kinect represents a potential source of difference across studies. Therefore, the general conclusions should be interpreted with caution.

Lastly, for reasons mentioned earlier, the authors did not examine validity, the degree to which a tool measures what it claims to measure. However, given the potential variety and the lack of any agreed-upon "gold standard" tool identified in the literature, a separate review is warranted to address validity. Reliability should always be interpreted with validity in mind to provide a complete assessment of the clinical appropriateness of a measuring tool.

Future Directions

Future reliability studies should focus on improving study design, with larger sample sizes (>80 participants) [138] and set recommended time intervals (2–7 days) between repeated measurements to increase confidence with results. Moreover, further investigations should report on absolute measures of reliability or measurement error to improve the overall risk of bias.

5. Conclusions

The primary result of our systematic review is that the Kinect and ambulatory sensor-based human motion tracking devices demonstrate moderate to good levels of intra- and inter-rater reliability to measure shoulder ROM. The assessment of reliability is an initial step in recommending a measuring tool for clinical use. Future research including the Kinect and other devices should investigate validity in well-designed, high-quality studies.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/s21248186/s1. Supplementary S1: Full search strategy.

Author Contributions: Conceptualization, P.B., D.B.A. and W.R.W.; methodology, P.B. and D.B.A.; formal analysis, P.B., M.P.; investigation, P.B. and D.B.A.; writing—original draft preparation, P.B.; writing—review and editing, P.B., D.B.A., M.P. and W.R.W.; supervision, M.P. and W.R.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this study are openly available in [96,97,102,106].

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Warner, J.J. Frozen Shoulder: Diagnosis and Management. *J. Am. Acad. Orthop. Surg.* **1997**, *5*, 130–140. [CrossRef] [PubMed]
2. Valentine, R.E.; Lewis, J.S. Intraobserver reliability of 4 physiologic movements of the shoulder in subjects with and without symptoms. *Arch. Phys. Med. Rehabil.* **2006**, *9*, 1242–1249. [CrossRef] [PubMed]
3. Green, S.; Buchbinder, R.; Glazier, R.; Forbes, A. Interventions for shoulder pain. *Cochrane Database Syst. Rev.* **2002**, *2*. [CrossRef]
4. Kapandji, I.A. *Physiology of the Joints*, 5th ed.; Upper Limb; Churchill Livingston: Oxford, UK, 1982; Volume 1.
5. Wu, G.; van der Helm, F.C.; Veeger, H.E.; Makhsous, M.; Van Roy, P.; Anglin, C.; Nagels, J.; Karduna, A.R.; McQuade, K.; Wang, X.; et al. International Society of Biomechanics. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: Shoulder, elbow, wrist and hand. *J. Biomech.* **2005**, *38*, 981–992. [CrossRef] [PubMed]
6. Tondu, B. Estimating Shoulder-Complex Mobility. *Appl. Bionics Biomech.* **2007**, *4*, 19–29. [CrossRef]
7. Muir, S.W.; Corea, C.L.; Beaupre, L. Evaluating change in clinical status: Reliability and measures of agreement for the assessment of glenohumeral range of motion. *N. Am. J. Sports Phys. Ther.* **2010**, *5*, 98–110. [PubMed]
8. Pearl, M.L.; Jackins, S.; Lippitt, S.B.; Sidles, J.A.; Matsen, F.A., 3rd. Humeroscapular positions in a shoulder range-of-motion-examination. *J. Shoulder Elb. Surg.* **1992**, *6*, 296–305. [CrossRef]
9. dos Santos, C.M.; Ferreira, G.; Malacco, P.L.; Sabino, G.S.; Moraes, G.F.D.S.; Felicio, D.C. Intra and inter examiner reliability and measurement error of goniometer and digital inclinometer use. *Rev. Bras. Med. Esporte* **2012**, *18*, 38–41.
10. American Academy of Orthopaedic Surgeons. *Joint Motion: Method of Measuring and Recording*; AAOS: Chicago, IL, USA, 1965.
11. Barnes, C.J.; Van Steyn, S.J.; Fischer, R.A. The effects of age, sex, and shoulder dominance on range of motion of the shoulder. *J. Shoulder Elb. Surg.* **2001**, *3*, 242–246. [CrossRef]
12. Roldán-Jiménez, C.; Cuesta-Vargas, A.I. Age-related changes analyzing shoulder kinematics by means of inertial sensors. *Clin. Biomech.* **2016**, *37*, 70–76. [CrossRef] [PubMed]
13. Gill, T.K.; Shanahan, E.M.; Tucker, G.R.; Buchbinder, R.; Hill, C.L. Shoulder range of movement in the general population: Age and gender stratified normative data using a community-based cohort. *BMC Musculoskelet. Disord.* **2020**, *21*, 676. [CrossRef] [PubMed]
14. Walston, J.D. Sarcopenia in older adults. *Curr. Opin. Rheumatol.* **2012**, *6*, 623–627. [CrossRef] [PubMed]
15. Soucie, J.M.; Wang, C.; Forsyth, A.; Funk, S.; Denny, M.; Roach, K.E.; Boone, D. Hemophilia Treatment Center Network. Range of motion measurements: Reference values and a database for comparison studies. *Haemophilia* **2011**, *3*, 500–507. [CrossRef] [PubMed]
16. Hayes, K.; Walton, J.R.; Szomor, Z.L.; Murrell, G.A.C. Reliability of five methods for assessing shoulder range of motion. *Aust. J. Physiother.* **2001**, *47*, 289–294. [CrossRef]
17. Riddle, D.L.; Rothstein, J.M.; Lamb, R.L. Goniometric reliability in a clinical setting. Shoulder measurements. *Phys. Ther.* **1987**, *67*, 668–673. [CrossRef]
18. Gajdosik, R.L.; Bohannon, R.W. Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. *Phys. Ther.* **1987**, *67*, 867–872. [CrossRef] [PubMed]
19. Laupattarakasem, W.; Sirichativapee, W.; Kowsuwon, W.; Sribunditkul, S.; Suibnugarn, C. Axial rotation gravity goniometer. A simple design of instrument and a controlled reliability study. *Clin. Orthop. Relat. Res.* **1990**, *251*, 271–274.
20. Boone, D.C.; Azen, S.P.; Lin, C.M.; Spence, C.; Baron, C.; Lee, L. Reliability of goniometric measurements. *Phys. Ther.* **1987**, *58*, 1355–1360. [CrossRef]
21. Pandya, S.; Florence, J.M.; King, W.M.; Robison, J.D.; Oxman, M.; Province, M.A. Reliability of goniometric measurements in patients with Duchenne muscular dystrophy. *Phys. Ther.* **1985**, *65*, 1339–1342. [CrossRef]
22. McVeigh, K.H.; Murray, P.M.; Heckman, M.G.; Rawal, B.; Peterson, J.J. Accuracy and Validity of Goniometer and Visual Assessments of Angular Joint Positions of the Hand and Wrist. *J. Hand Surg. Am.* **2016**, *41*, e21–e35. [CrossRef] [PubMed]
23. Bovens, A.M.; van Baak, M.A.; Vrencken, J.G.; Wijnen, J.A.; Verstappen, F.T. Variability and reliability of joint measurements. *Am. J. Sports Med.* **1990**, *18*, 58–63. [CrossRef]
24. Walker, J.M.; Sue, D.; Miles-Elkousy, N.; Ford, G.; Trevelyan, H. Active mobility of the extremities in older subjects. *Phys. Ther.* **1984**, *64*, 919–923. [CrossRef] [PubMed]
25. Goodwin, J.; Clark, C.; Deakes, J.; Burdon, D.; Lawrence, C. Clinical methods of goniometry: A comparative study. *Disabil. Rehabil.* **1992**, *14*, 10–15. [CrossRef]
26. Lea, R.D.; Gerhardt, J.J. Range-of-motion measurements. *J. Bone Jt. Surg. Am.* **1995**, *77*, 784–798. [CrossRef] [PubMed]
27. Vafadar, A.K.; Côté, J.N.; Archambault, P.S. Interrater and Intrarater Reliability and Validity of 3 Measurement Methods for Shoulder-Position Sense. *J. Sport Rehabil.* **2016**, *25*. [CrossRef] [PubMed]
28. Keogh, J.W.L.; Espinosa, H.G.; Grigg, J. Evolution of smart devices and human movement apps: Recommendations for use in sports science education and practice. *J. Fit. Res.* **2016**, *5*, 14–15.

29. Mosa, A.S.M.; Yoo, I.; Sheets, L. A Systematic Review of Healthcare Applications for Smartphones. *BMC. Med. Inform. Decis. Mak.* **2012**, *12*, 67. [CrossRef]
30. Bruyneel, A.V. Smartphone Applications for Range of Motion Measurement in Clinical Practice: A Systematic Review. *Stud. Health Technol. Inform.* **2020**, *270*, 1389–1390. [CrossRef] [PubMed]
31. Vercelli, S.; Sartorio, F.; Bravini, E.; Ferriero, G. DrGoniometer: A reliable smartphone app for joint angle measurement. *Br. J. Sports Med.* **2017**, *51*, 1703–1704. [CrossRef]
32. Kolber, M.J.; Hanney, W.J. The reliability and concurrent validity of shoulder mobility measurements using a digital inclinometer and goniometer: A technical report. *Int. J. Sports Phys. Ther.* **2012**, *7*, 306–313. [PubMed]
33. Rodgers, M.M.; Alon, G.; Pai, V.M.; Conroy, R.S. Wearable technologies for active living and rehabilitation: Current research challenges and future opportunities. *J. Rehabil. Assist. Technol. Eng.* **2019**, *6*, 2055668319839607. [CrossRef]
34. Mourcou, Q.; Fleury, A.; Diot, B.; Franco, C.; Vuillerme, N. Mobile Phone-Based Joint Angle Measurement for Functional Assessment and Rehabilitation of Proprioception. *Biomed. Res. Int.* **2015**, *2015*, 328142. [CrossRef]
35. Faisal, A.I.; Majumder, S.; Mondal, T.; Cowan, D.; Naseh, S.; Deen, M.J. Monitoring Methods of Human Body Joints: State-of-the-Art and Research Challenges. *Sensors* **2019**, *19*, 2629. [CrossRef]
36. TechTarget. Kinect. Available online: http://searchhealthit.techtarget.com/definition/Kinect (accessed on 14 October 2021).
37. Anton, D.; Berges, I.; Bermúdez, J.; Goñi, A.; Illarramendi, A. A Telerehabilitation System for the Selection, Evaluation and Remote Management of Therapies. *Sensors* **2018**, *18*, 1459. [CrossRef]
38. Antón, D.; Nelson, M.; Russell, T.; Goñi, A.; Illarramendi, A. Validation of a Kinect-based telerehabilitation system with total hip replacement patients. *J. Telemed. Telecare* **2016**, *22*, 192–197. [CrossRef]
39. Kang, Y.S.; Chang, Y.J.; Howell, S.R. Using a kinect-based game to teach oral hygiene in four elementary students with intellectual disabilities. *J. Appl. Res. Intellect. Disabil.* **2021**, *34*, 606–614. [CrossRef]
40. Boutsika, E. Kinect in Education: A Proposal for Children with Autism. *Procedia Comput. Sci.* **2014**, *27*, 123–129. [CrossRef]
41. Pfister, A.; West, A.M.; Bronner, S.; Noah, J.A. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *J. Med. Eng. Technol.* **2014**, *38*, 274–280. [CrossRef]
42. Mangal, N.K.; Tiwari, A.K. Kinect v2 tracked Body Joint Smoothing for Kinematic Analysis in Musculoskeletal Disorders. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 5769–5772. [CrossRef]
43. Bonnechère, B.; Sholukha, V.; Omelina, L.; Van Sint Jan, S.; Jansen, B. 3D Analysis of Upper Limbs Motion during Rehabilitation Exercises Using the Kinect Sensor: Development, Laboratory Validation and Clinical Application. *Sensors* **2018**, *18*, 2216. [CrossRef]
44. Moon, S.; Park, Y.; Ko, D.W.; Suh, I.H. Multiple Kinect Sensor Fusion for Human Skeleton Tracking Using Kalman Filtering. *Int. J. Adv. Robot. Syst.* **2016**, *13*, 65. [CrossRef]
45. Wiedemann, L.G.; Planinc, R.; Nemec, I.; Kampel, M. Performance evaluation of joint angles obtained by the Kinect V2. In Proceedings of the IET International Conference on Technologies for Active and Assisted Living (TechAAL), London, UK, 5 November 2015; pp. 1–6. [CrossRef]
46. Zerpa, C.; Lees, C.; Patel, P.; Pryzsucha, E. The Use of Microsoft Kinect for Human Movement Analysis. *Int. J. Sports Sci.* **2015**, *5*, 120–127. [CrossRef]
47. Zhang, Z.; Ji, L.; Huang, Z.; Wu, J. Adaptive Information Fusion for Human Upper Limb Movement Estimation. *IEEE Trans. Syst. Man. Cybern. Part A Syst. Hum.* **2012**, *42*, 1100–1108. [CrossRef]
48. Saber-Sheikh, K.; Bryant, E.C.; Glazzard, C.; Hamel, A.; Lee, R.Y. Feasibility of using inertial sensors to assess human movement. *Man. Ther.* **2010**, *15*, 122–125. [CrossRef]
49. Carnevale, A.; Longo, U.G.; Schena, E.; Massaroni, C.; Lo Presti, D.; Berton, A.; Candela, V.; Denaro, V. Wearable systems for shoulder kinematics assessment: A systematic review. *BMC Musculoskelet. Disord.* **2019**, *20*, 546. [CrossRef] [PubMed]
50. Robert-Lachaine, X.; Mecheri, H.; Muller, A.; Larue, C.; Plamondon, A. Validation of a low-cost inertial motion capture system for whole-body motion analysis. *J. Biomech.* **2020**, *99*, 109520. [CrossRef]
51. Bewes, R.; Callaway, A.; Williams, J. Inertial sensor measurement of shoulder joint position sense: Reliability and consistency. *Physiotherapy* **2020**, *107*, e114. [CrossRef]
52. Walmsley, C.P.; Williams, S.A.; Grisbrook, T.; Elliott, C.; Imms, C.; Campbell, A. Measurement of Upper Limb Range of Motion Using Wearable Sensors: A Systematic Review. *Sports Med. Open* **2018**, *4*, 53. [CrossRef] [PubMed]
53. Garimella, R.; Peeters, T.; Beyers, K.; Truijen, S.; Huysmans, T.; Verwulgen, S. Capturing Joint Angles of the Off-site Human Body. In Proceedings of the IEEE Sensors, New Delhi, India, 28–31 October 2018; pp. 1–4.
54. Kuhlmann, T.; Garaizar, P.; Reips, U.D. Smartphone sensor accuracy varies from device to device in mobile research: The case of spatial orientation. *Behav. Res.* **2021**, *53*, 22–33. [CrossRef] [PubMed]
55. Hsu, Y.; Wang, J.; Lin, Y.; Chen, S.; Tsai, Y.; Chu, C.; Chang, C. A wearable inertial-sensing-based body sensor network for shoulder range of motion assessment. In Proceedings of the 2013 1st International Conference on Orange Technologies (ICOT), Tainan, Taiwan, 12–16 March 2013; pp. 328–331. [CrossRef]
56. Streiner, D.; Norman, G. *Health Measurement Scales: A Practical Guide to Their Development and Use*, 3rd ed.; Oxford University Press: New York, NY, USA, 2003.

57. Scano, A.; Mira, R.M.; Cerveri, P.; Molinari Tosatti, L.; Sacco, M. Analysis of Upper-Limb and Trunk Kinematic Variability: Accuracy and Reliability of an RGB-D Sensor. *Multimodal Technol. Interact.* **2020**, *4*, 14. [CrossRef]
58. López, N.; Perez, E.; Tello, E.; Rodrigo, A.; Valentinuzzi, M.E. Statistical Validation for Clinical Measures: Repeatability and Agreement of Kinect-Based Software. *BioMed Res. Int.* **2018**, *2018*, 6710595. [CrossRef] [PubMed]
59. Wilson, J.D.; Khan-Perez, J.; Marley, D.; Buttress, S.; Walton, M.; Li, B.; Roy, B. Can shoulder range of movement be measured accurately using the Microsoft Kinect sensor plus Medical Interactive Recovery Assistant (MIRA) software? *J. Shouder Elb. Surg.* **2017**, *26*, e382–e389. [CrossRef] [PubMed]
60. Lee, S.H.; Yoon, C.; Chung, S.G.; Kim, H.C.; Kwak, Y.; Park, H.; Kim, K. Measurement of shoulder range of motion in patients with adhesive capsulitis using a kinect. *PLoS ONE* **2015**, *10*, e0129398. [CrossRef] [PubMed]
61. Poitras, I.; Bielmann, M.; Campeau-Lecours, A.; Mercier, C.; Bouyer, L.J.; Roy, J.S. Validity of Wearable Sensors at the Shoulder Joint: Combining Wireless Electromyography Sensors and Inertial Measurement Units to Perform Physical Workplace Assessments. *Sensors* **2019**, *19*, 1885. [CrossRef] [PubMed]
62. Morrow, M.M.B.; Lowndes, B.; Fortune, E.; Kaufman, K.R.; Hallbeck, M.S. Validation of Inertial Measurement Units for Upper Body Kinematics. *J. Appl. Biomech.* **2017**, *33*, 227–232. [CrossRef]
63. Lin, Y.C.; Tsai, Y.J.; Hsu, Y.L.; Yen, M.H.; Wang, J.S. Assessment of Shoulder Range of Motion Using a Wearable Inertial Sensor Network. *IEEE Sens. J.* **2021**, *21*, 15330–15341. [CrossRef]
64. Carey, M.A.; Laird, D.E.; Murray, K.A.; Stevenson, J.R. Reliability, validity, and clinical usability of a digital goniometer. *Work* **2010**, *36*, 55–66. [CrossRef]
65. Hannah, D.C.; Scibek, J.S. Collecting shoulder kinematics with electromagnetic tracking systems and digital inclinometers: A review. *World J. Orthop.* **2015**, *6*, 783–794. [CrossRef] [PubMed]
66. Scibek, J.S.; Carcia, C.R. Assessment of scapulohumeral rhythm for scapular plane shoulder elevation using a modified digital inclinometer. *World J. Orthop.* **2012**, *3*, 87–94. [CrossRef] [PubMed]
67. Johnson, M.P.; McClure, P.W.; Karduna, A.R. New method to assess scapular upward rotation in subjects with shoulder pathology. *J. Orthop. Sports Phys. Ther.* **2001**, *31*, 81–89. [CrossRef] [PubMed]
68. Pichonnaz, C.; Aminian, K.; Ancey, C.; Jaccard, H.; Lécureux, E.; Duc, C.; Farron, A.; Jolles, B.M.; Gleeson, N. Heightened clinical utility of smartphone versus body-worn inertial system for shoulder function B-B score. *PLoS ONE* **2017**, *12*, e0174365. [CrossRef] [PubMed]
69. Johnson, L.B.; Sumner, S.; Duong, T.; Yan, P.; Bajcsy, R.; Abresch, R.T.; de Bie, E.; Han, J.J. Validity and reliability of smartphone magnetometer-based goniometer evaluation of shoulder abduction—A pilot study. *Man. Ther.* **2015**, *20*, 777–782. [CrossRef]
70. Keogh, J.W.L.; Cox, A.; Anderson, S.; Liew, B.; Olsen, A.; Schram, B.; Furness, J. Reliability and validity of clinically accessible smartphone applications to measure joint range of motion: A systematic review. *PLoS ONE* **2019**, *14*, e0215806. [CrossRef]
71. McDowell, I. *Measuring Health: A Guide to Rating Scales and Questionnaires*, 3rd ed.; Oxford University Press: New York, NY, USA, 2006.
72. Norkin, C.C.; White, D.J.; Torres, J.; Greene, M.J.; Grochowska, L.L.; Malone, T.W. *Measurement of Joint Motion: A Guide to Goniometry*; F.A. Davis: Philadelphia, PA, USA, 1995.
73. Altman, D.G.; Bland, J.M. Measurement in medicine: The analysis of method comparison studies. *Statistician* **1983**, *32*, 307–317. [CrossRef]
74. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *1*, 307–310. [CrossRef]
75. de Vet, H.C.; Terwee, C.B.; Knol, D.L.; Bouter, L.M. When to use agreement versus reliability measures. *J. Clin. Epidemiol.* **2006**, *10*, 1033–1039. [CrossRef] [PubMed]
76. Stratford, P.W. Getting more from the literature: Estimating the standard error of measurement from reliability studies. *Physiother. Can.* **2004**, *56*, 27–30. [CrossRef]
77. Webster, D.; Celik, O. Systematic review of Kinect applications in elderly care and stroke rehabilitation. *J. Neuroeng. Rehabil.* **2014**, *11*, 108. [CrossRef]
78. Xavier-Rocha, T.B.; Carneiro, L.; Martins, G.C.; Vilela-JÚnior, G.B.; Passos, R.P.; Pupe, C.C.B.; Nascimento, O.J.M.D.; Haikal, D.S.; Monteiro-Junior, R.S. The Xbox/Kinect use in poststroke rehabilitation settings: A systematic review. *Arq. Neuropsiquiatr.* **2020**, *78*, 361–369. [CrossRef]
79. Knippenberg, E.; Verbrugghe, J.; Lamers, I.; Palmaers, S.; Timmermans, A.; Spooren, A. Markerless motion capture systems as training device in neurological rehabilitation: A systematic review of their use, application, target population and efficacy. *J. Neuroeng. Rehabil.* **2017**, *14*, 61. [CrossRef]
80. Garcia-Agundez, A.; Folkerts, A.K.; Konrad, R.; Caserman, P.; Tregel, T.; Goosses, M.; Göbel, S.; Kalbe, E. Recent advances in rehabilitation for Parkinson's Disease with Exergames: A Systematic Review. *J. Neuroeng. Rehabil.* **2019**, *16*, 17. [CrossRef]
81. Marotta, N.; Demeco, A.; Indino, A.; de Scorpio, G.; Moggio, L.; Ammendolia, A. Nintendo Wii versus Xbox Kinect for functional locomotion in people with Parkinson's disease: A systematic review and network meta-analysis. *Disabil. Rehabil.* **2020**, 1–6. [CrossRef]
82. Barry, G.; Galna, B.; Rochester, L. The role of exergaming in Parkinson's disease rehabilitation: A systematic review of the evidence. *J. NeuroEng. Rehabil.* **2014**, *11*, 33. [CrossRef] [PubMed]

83. Puh, U.; Hoehlein, B.; Deutsch, J.E. Validity and Reliability of the Kinect for Assessment of Standardized Transitional Movements and Balance: Systematic Review and Translation into Practice. *Phys. Med. Rehabil. Clin. N. Am.* **2019**, *30*, 399–422. [CrossRef] [PubMed]
84. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.C.; Ioannidis, J.P.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med.* **2009**, *6*, e1000100. [CrossRef] [PubMed]
85. van de Pol, R.J.; van Trijffel, E.; Lucas, C. Inter-rater reliability for measurement of passive physiological range of motion of upper extremity joints is better if instruments are used: A systematic review. *J. Physiother.* **2010**, *56*, 7–17. [CrossRef]
86. Denteneer, L.; Stassijns, G.; De Hertogh, W.; Truijen, S.; Van Daele, U. Inter- and Intrarater Reliability of Clinical Tests Associated with Functional Lumbar Segmental Instability and Motor Control Impairment in Patients with Low Back Pain: A Systematic Review. *Arch. Phys. Med. Rehabil.* **2017**, *98*, 151–164.e6. [CrossRef] [PubMed]
87. Swinkels, R.; Bouter, L.; Oostendorp, R.; van den Ende, C. Impairment measures in rheumatic disorders for rehabilitation medicine and allied health care: A systematic review. *Rheumatol. Int.* **2005**, *25*, 501–512. [CrossRef]
88. Mokkink, L.B.; Boers, M.; van der Vleuten, C.P.M.; Bouter, L.M.; Alonso, J.; Patrick, D.L.; de Vet, H.C.W.; Terwee, C.B. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. *BMC Med. Res. Methodol.* **2020**, *20*, 293. [CrossRef] [PubMed]
89. Anderson, D.B.; Mathieson, S.; Eyles, J.; Maher, C.G.; Van Gelder, J.M.; Tomkins-Lane, C.C.; Ammendolia, C.; Bella, V.; Ferreira, M.L. Measurement properties of walking outcome measures for neurogenic claudication: A systematic review and meta-analysis. *Spine J.* **2019**, *19*, 1378–1396. [CrossRef]
90. Dobson, F.; Hinman, R.S.; Hall, M.; Terwee, C.B.; Roos, E.M.; Bennell, K.L. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: A systematic review. *Osteoarthr. Cartil.* **2012**, *12*, 1548–1562. [CrossRef] [PubMed]
91. Awan, R.; Smith, J.; Boon, A.J. Measuring shoulder internal rotation range of motion: A comparison of 3 techniques. *Arch. Phys. Med. Rehabil.* **2002**, *83*, 1229–1234. [CrossRef] [PubMed]
92. Beshara, P.; Chen, J.; Lagadec, P.; Walsh, W.R. Test-Retest and Intra-rater Reliability of Using Inertial Sensors and Its Integration with Microsoft Kinect™ to Measure Shoulder Range-of-Motion. In *Internet of Things Technologies for HealthCare*; Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering Series; Springer International Publishing: Cham, Switzerland, 2016; Volume 187.
93. Bonnechère, B.; Jansen, B.; Salvia, P.; Bouzahouene, H.; Omelina, L.; Moiseev, F.; Sholukha, V.; Cornelis, J.; Rooze, M.; Van Sint Jan, S. Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry. *Gait Posture* **2014**, *39*, 593–598. [CrossRef] [PubMed]
94. Cai, L.; Ma, Y.; Xiong, S.; Zhang, Y. Validity and Reliability of Upper Limb Functional Assessment Using the Microsoft Kinect V2 Sensor. *Appl. Bionics Biomech.* **2019**, *2019*, 175240. [CrossRef] [PubMed]
95. Chan, C.; Sihler, E.; Kijek, T.; Miller, B.S.; Hughes, R.E. A mobile computing tool for collecting clinical outcomes data from shoulder patients. *J. Musculoskelet. Res.* **2010**, *13*, 127–136. [CrossRef]
96. Chen, Y.P.; Lin, C.Y.; Tsai, M.J.; Chuang, T.Y.; Lee, O.K. Wearable Motion Sensor Device to Facilitate Rehabilitation in Patients with Shoulder Adhesive Capsulitis: Pilot Study to Assess Feasibility. *J. Med. Internet Res.* **2020**, *22*, e17032. [CrossRef] [PubMed]
97. Cools, A.M.; De Wilde, L.; Van Tongel, A.; Ceyssens, C.; Ryckewaert, R.; Cambier, D.C. Measuring shoulder external and internal rotation strength and range of motion: Comprehensive intra-rater and inter-rater reliability study of several testing protocols. *J. Shoulder Elb. Surg.* **2014**, *23*, 1454–1461. [CrossRef] [PubMed]
98. Correll, S.; Field, J.; Hutchinson, H.; Mickevicius, G.; Fitzsimmons, A.; Smoot, B. Reliability and Validity of the Halo Digital goniometer for shoulder range of motion in healthy subjects. *Int. J. Sports Phys. Ther.* **2018**, *13*, 707–714. [CrossRef] [PubMed]
99. Çubukçu, B.; Yüzgeç, U.; Zileli, R.; Zileli, A. Reliability and validity analyzes of Kinect V2 based measurement system for shoulder motions. *Med. Eng. Phys.* **2020**, *76*, 20–31. [CrossRef]
100. Cuesta-Vargas, A.I.; Roldán-Jiménez, C. Validity and reliability of arm abduction angle measured on smartphone: A cross-sectional study. *BMC Musculoskelet. Disord.* **2016**, *17*, 93. [CrossRef] [PubMed]
101. Da Cunha Neto, J.S.; Rebouças Filho, P.P.; Da Silva, G.P.F.; Da Cunha Olegario, N.B.; Duarte, J.B.F.; De Albuquerque, V.H.C. Dynamic Evaluation and Treatment of the Movement Amplitude Using Kinect Sensor. *IEEE Access* **2018**, *6*, 17292–17305. [CrossRef]
102. De Baets, L.; Vanbrabant, S.; Dierickx, C.; van der Straaten, R.; Timmermans, A. Assessment of Scapulothoracic, Glenohumeral, and Elbow Motion in Adhesive Capsulitis by Means of Inertial Sensor Technology: A Within-Session, Intra-Operator and Inter-Operator Reliability and Agreement Study. *Sensors* **2020**, *20*, 876. [CrossRef]
103. de Winter, A.F.; Heemskerk, M.A.; Terwee, C.B.; Jans, M.P.; Devillé, W.; van Schaardenburg, D.J.; Scholten, R.J.; Bouter, L.M. Inter-observer reproducibility of measurements of range of motion in patients with shoulder pain using a digital inclinometer. *BMC Musculoskelet. Disord.* **2004**, *5*, 18. [CrossRef]
104. Dougherty, J.; Walmsley, S.; Osmotherly, P.G. Passive range of movement of the shoulder: A standardized method for measurement and assessment of intrarater reliability. *J. Manip. Physiol. Ther.* **2015**, *38*, 218–224. [CrossRef] [PubMed]

105. Hawi, N.; Liodakis, E.; Musolli, D.; Suero, E.M.; Stuebig, T.; Claassen, L.; Kleiner, C.; Krettek, C.; Ahlers, V.; Citak, M. Range of motion assessment of the shoulder and elbow joints using a motion sensing input device: A pilot study. *Technol. Health Care* **2014**, *22*, 289–295. [CrossRef]
106. Huber, M.E.; Seitz, A.L.; Leeser, M.; Sternad, D. Validity and reliability of Kinect skeleton for measuring shoulder joint angles: A feasibility study. *Physiotherapy* **2015**, *101*, 389–393. [CrossRef]
107. Hwang, S.; Tsai, C.Y.; Koontz, A.M. Feasibility study of using a Microsoft Kinect for virtual coaching of wheelchair transfer techniques. *Biomed. Tech.* **2017**, *62*, 307–313. [CrossRef] [PubMed]
108. Kolber, M.J.; Vega, F.; Widmayer, K.; Cheng, M.S. The reliability and minimal detectable change of shoulder mobility measurements using a digital inclinometer. *Physiother. Theory Pract.* **2011**, *27*, 176–184. [CrossRef]
109. Kolber, M.J.; Fuller, C.; Marshall, J.; Wright, A.; Hanney, W.J. The reliability and concurrent validity of scapular plane shoulder elevation measurements using a digital inclinometer and goniometer. *Physiother. Theory Pract.* **2012**, *28*, 161–168. [CrossRef]
110. Lim, J.Y.; Kim, T.H.; Lee, J.S. Reliability of measuring the passive range of shoulder horizontal adduction using a smartphone in the supine versus the side-lying position. *J. Phys. Ther. Sci.* **2015**, *10*, 3119–3122. [CrossRef] [PubMed]
111. Mejia-Hernandez, K.; Chang, A.; Eardley-Harris, N.; Jaarsma, R.; Gill, T.K.; McLean, J.M. Smartphone applications for the evaluation of pathologic shoulder range of motion and shoulder scores-a comparative study. *JSES Open Access* **2018**, *2*, 109–114. [CrossRef] [PubMed]
112. Milgrom, R.; Foreman, M.; Standeven, J.; Engsberg, J.R.; Morgan, K.A. Reliability and validity of the Microsoft Kinect for assessment of manual wheelchair propulsion. *J. Rehabil. Res. Dev.* **2016**, *53*, 901–918. [CrossRef]
113. Mitchell, K.; Gutierrez, S.B.; Sutton, S.; Morton, S.; Morgenthaler, A. Reliability and validity of goniometric iPhone applications for the assessment of active shoulder external rotation. *Physiother. Theory Pract.* **2014**, *30*, 521–525. [CrossRef]
114. Picerno, P.; Viero, V.; Donati, M.; Triossi, T.; Tancredi, V.; Melchiorri, G. Ambulatory assessment of shoulder abduction strength curve using a single wearable inertial sensor. *J. Rehabil. Res. Dev.* **2015**, *52*, 171–180. [CrossRef]
115. Poser, A.; Ballarin, R.; Piu, P.; Venturin, D.; Plebani, G.; Rossi, A. Intra and inter examiner reliability of the range of motion of the shoulder in asymptomatic subjects, by means of digital inclinometers. *Sci. Riabil.* **2015**, *17*, 17–25.
116. Ramos, M.M.; Carnaz, L.; Mattiello, S.M.; Karduna, A.R.; Zanca, G.G. Shoulder and elbow joint position sense assessment using a mobile app in subjects with and without shoulder pain-between-days reliability. *Phys. Ther. Sport* **2019**, *37*, 157–163. [CrossRef]
117. Rigoni, M.; Gill, S.; Babazadeh, S.; Elsewaisy, O.; Gillies, H.; Nguyen, N.; Pathirana, P.N.; Page, R. Assessment of Shoulder Range of Motion Using a Wireless Inertial Motion Capture Device—A Validation Study. *Sensors* **2019**, *19*, 1781. [CrossRef]
118. Schiefer, C.; Kraus, T.; Ellegast, R.P.; Ochsmann, E. A technical support tool for joint range of motion determination in functional diagnostics—An inter-rater study. *J. Occup. Med. Toxicol.* **2015**, *10*, 16. [CrossRef]
119. Scibek, J.S.; Carcia, C.R. Validation and repeatability of a shoulder biomechanics data collection methodology and instrumentation. *J. Appl. Biomech.* **2013**, *29*, 609–615. [CrossRef] [PubMed]
120. Shin, S.H.; Lee, O.S.; Oh, J.H.; Kim, S.H. Within-day reliability of shoulder range of motion measurement with a smartphone. *Man. Ther.* **2012**, *17*, 298–304. [CrossRef]
121. Walker, H.; Pizzari, T.; Wajswelner, H.; Blanch, P.; Schwab, L.; Bennell, K.; Gabbe, B. The reliability of shoulder range of motion measures in competitive swimmers. *Phys. Ther. Sport* **2016**, *21*, 26–30. [CrossRef]
122. Werner, B.C.; Holzgrefe, R.E.; Griffin, J.W.; Lyons, M.L.; Cosgrove, C.T.; Hart, J.M.; Brockmeier, S.F. Validation of an innovative method of shoulder range-of-motion measurement using a smartphone clinometer application. *J. Shoulder Elb. Surg.* **2014**, *23*, e275–e282. [CrossRef] [PubMed]
123. Frost, M.H.; Reeve, B.B.; Liepa, A.M.; Stauffer, J.W.; Hays, R.D. Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health* **2007**, S94–S105. [CrossRef] [PubMed]
124. Fayers, P.; Machin, D. *Quality of Life. Assessment, Analysis and Interpretation*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2007.
125. Sim, J.; Wright, C.C. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys. Ther.* **2005**, *85*, 257–268. [CrossRef] [PubMed]
126. Monnier, A.; Heuer, J.; Norman, K.; Ang, B.O. Inter- and intra-observer reliability of clinical movement-control tests for marines. *BMC Musculoskelet. Disord.* **2012**, *13*, 263. [CrossRef] [PubMed]
127. Tegern, M.; Aasa, U.; Äng, B.O.; Harms-Ringdahl, K.; Larsson, H. Inter-rater and test-retest reliability of movement control tests for the neck, shoulder, thoracic, lumbar, and hip regions in military personnel. *PLoS ONE* **2018**, *13*, e0204552. [CrossRef]
128. Nayak, B.K. Understanding the relevance of sample size calculation. *Indian J. Ophthalmol.* **2010**, *58*, 469–470. [CrossRef]
129. Mokkink, L.B.; de Vet, H.C.W.; Prinsen, C.A.C.; Patrick, D.L.; Alsonso, J.; Bouter, L.M.; Terwee, C.B. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual. Life Res.* **2018**, *27*, 1171–1179. [CrossRef]
130. Carter, R.E.; Lubinsky, J.; Domholdt, E. *Rehabilitation Research: Principles and Applications*, 4th ed.; Elsevier Inc.: St. Louis, MO, USA, 2011.
131. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
132. Hobart, J.C.; Cano, S.J.; Warner, T.T.; Thompson, A.J. What sample sizes for reliability and validity studies in neurology? *J. Neurol.* **2012**, *259*, 2681–2694. [CrossRef]

133. Wright, J.G.; Feinstein, A.R. Improving the reliability of orthopaedic measurements. *J. Bone Jt. Surg. Br.* **1992**, *74*, 287–291. [CrossRef]
134. Sabari, J.S.; Maltzev, I.; Lubarsky, D.; Liszkay, E.; Homel, P. Goniometric assessment of shoulder range of motion: Comparison of testing in supine and sitting positions. *Arch. Phys. Med. Rehabil.* **1998**, *6*, 64751. [CrossRef]
135. Kim, B.R.; Yi, D.; Yim, J. Effect of postural change on shoulder joint internal and external rotation range of motion in healthy adults in their 20s. *Phys. Ther. Rehabil. Sci.* **2019**, *8*, 152–157. [CrossRef]
136. Xu, X.; McGorry, R.W. The validity of the first and second generation Microsoft Kinect for identifying joint center locations during static postures. *Appl. Ergon.* **2015**, *49*, 47–54. [CrossRef] [PubMed]
137. Prinsen, C.A.C.; Mokkink, L.B.; Bouter, L.M.; Alonso, J.; Patrick, D.L.; de Vet, H.C.W.; Terwee, C.B. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual. Life Res.* **2018**, *27*, 1147–1157. [CrossRef] [PubMed]
138. Bujang, M.A.; Baharum, N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Arch. Orofac. Sci.* **2017**, *12*, 1–11.