Assessing reliability of movement screenings using a markerless motion capture system

Jonathan D. Hauenstein[a], Alan Huebner[a], John P. Wagle[b], Emma R. Cobian[a], Joseph Cummings[a], Caroline Hills[a], Megan McGinty[a], Mandy Merritt[b], Sam Rosengarten[a,c], Kyle Skinner[b], Michael Szemborski[b], Leigh Wojtkiewicz[d]

[a]University of Notre Dame, Department of Applied and Computational Mathematics and Statistics, 102G Crowley Hall, Notre Dame, IN, USA

[b]University of Notre Dame, Sports Performance, 113 Joyce Center, Notre Dame, IN 46556, USA

[c]Baltimore Ravens, UnderArmour Performance Center, 1 Winning Drive, Owings Mills, MD 21117, USA

[d]University of Notre Dame, Data & Analytics, 113 Joyce Center, Notre Dame, IN 46556, USA

Corresponding Author:
Jonathan D. Hauenstein, 102G Crowley Hall, Notre Dame, IN 46556, USA, hauenstein@nd.edu

Abstract

Background: Movement screenings are commonly used to detect unfavorable movement patterns. Due to the burdensomeness of markers, markerless motion capture systems have been developed to track three-dimensional motion.

Purpose: To determine the reliability of movement screenings assessed using a markerless motion capture system when comparing the results of multiple systems and multiple collection periods across 39 kinematic variables.

Study Design: Cross-Sectional Study.

Methods: This research study investigated the inter- and intra-session reliability of a commercially available markerless motion capture system. A total of 39 kinematic variables were considered arising from 10 fundamental upper and lower body movements that are typical of a screening procedure in sports performance using 21 recreationally active participants between the ages of 18 and 22. The data was statistically analyzed in terms of relative error via the intraclass correlation coefficient and absolute error via the typical error.

Results: The results show that both the inter- and intra-rater reliability as measured by the intraclass correlation coefficient was at least moderate with the vast majority having excellent reliability with regards to all movements and corresponding variables. Although maximum knee valgus angles were the kinematic variables with lowest inter-rater reliability with moderate relative agreement, there was agreement in absolute terms with a typical error of less than 1.3 degrees.

Conclusion: Markerless motion capture is reliable for assessing movement screenings when comparing the results of multiple systems or multiple collection periods. However, practitioners should consider both relative and absolute agreements when applying information provided by these systems. Ultimately, markerless motion capture appears to provide reliable measurements of joint position during a movement screen, which allows for more objective evaluation of the direction and subsequent success of interventions.

Clinical Relevance: Markerless motion capture systems may assist clinicians by reliably assessing movement screenings using different systems over different collection periods.

What is known about this subject: Movement screening assessments are routinely integrated into the support of athletes most commonly as a precautionary measure for detecting potentially unfavorable movement patterns that may put an athlete at greater than normal injury risk or be problematic in the execution of technical sporting actions (e.g., throwing). Since marker-based systems are burdensome, markerless systems provide an easier to use alternative for clinicians but there is a paucity of literature that assesses the reliability of markerless systems.

What this study adds to existing knowledge: This research study indicates that markerless motion capture is reliable when comparing the results of multiple systems or multiple collection periods. Such systems provide a suitable, even superior, alternative to subjectively scored screening protocols, an advantage brought about by the granularity afforded in a markerless motion capture system. Reliably describing joint position during a movement screen allows for more objective evaluation of the direction and subsequent success of interventions. Undoubtedly, markerless motion capture will continue to scale its presence in sports medicine.

Keywords

movement screening, markerless motion capture, inter-rater reliability, intra-rater reliability, kinematic variables

1. Introduction

Movement screening assessments are routinely integrated into the support of athletes most commonly as a precautionary measure to identify several abnormal patterns of motion which adversely affect their performance in sport and predispose them to injury. The screens are typically best suited as a complimentary piece to an objective clinical assessment by a trained sports medicine practitioner, largely because of the subjective nature of scoring and the inherent inter- and intra-rater reliability issues.[26] Removing rating subjectivity of movement screens may make these data more valuable and equip practitioners with objective information from both their clinical and functional assessments.

Collection of kinematic data during a movement screen or sporting action is hardly novel, but the typical methods for obtaining this data require optical tracking of reflective markers.[25] The reflective markers are burdensome during the data acquisition process and force environmental constraints that reduce the construct validity of the assessment itself. Hence, the development of markerless motion capture systems to track three-dimensional motion resolves the need for markers. This affords collection of more simplistic biomechanical athletic movements (e.g., movement screening) and has obvious potential to capture the complexity of the sporting action itself within its native context (e.g., in-competition). However, there exists a paucity of literature to date that assesses the reliability of these systems and their underlying algorithmic models[6,12,23] and, furthermore, the existing research compares to marker-based systems.[5,9,18] This is logical considering marker-based systems are the current benchmark, but the models used to drive these systems are not without their own limitations (e.g., occlusion, skin artifacts) that may contribute to a biased interpretation when compared to markerless systems.[4,11] Similarly, two-dimensional markerless measurements of movement are prevalent, especially due to the increased processing capabilities of smartphones and tablets. These options have limitations with respect to simplicity (i.e., two-dimensional vs. three-dimensional modeling) that may limit practical utility and validity. Moreover, there is low agreement relative to marker-based systems with error magnitudes similar to manual techniques such as a goniometry.[1,22] Weighing these limitations collectively, it may be more appropriate to establish repeatability within an individual markerless system and between two independent markerless systems. To date, no such investigation exists, leaving a meaningful gap in the motion capture literature. Therefore, the purpose of the current study is to investigate the inter- and intra-session reliability of a commercially available markerless motion capture system and the respective algorithms used to derive whole body and joint-specific kinematic variables. In particular, this study considered testing the hypothesis that such a markerless motion capture system and algorithms would be reliable under both inter- and intra-session conditions for whole body and joint-specific measurements typical of a screening procedure used in sports performance.

2. Methods

Twenty-one recreationally active participants (6 male, 15 female) volunteered to participate in the study (mean ± sd: height = 174.53 ± 9.64 cm, weight = 68.64 ± 11.24 kg, age between 18 and 22 years old). As described below, reliability was assessed primarily using the intraclass correlation coefficient (ICC), thus sample size planning was conducted in reference to this statistic. Specifically, the method provided by Bonett[3] was followed to obtain a sample size so that reasonably narrow confidence intervals for the ICC could be obtained. Due to the statistical properties of the ICC, a higher ICC requires fewer observations to obtain a narrow confidence interval (CI). Thus, it was determined that a sample size of approximately 20 subjects was adequate to identify the variables that were measured with high reliability.

Participants signed written informed consent forms approved by the University of Notre Dame Institutional Review Board and attested that they have no current injuries that would influence their ability to perform the protocol. Following a self-selected warm-up lasting approximately three minutes, participants were asked to perform two sets of three repetitions of each movement, except for the single leg hops which had five jumps each, within a series of fundamental upper and lower body movements that are typical of a screening procedure in sports performance (Table 1). Each movement was separated by approximately 30 seconds of standing, passive rest where study staff shared standardized coaching instructions describing the subsequent movement. Each set was separated by 5-10 minutes of seated, passive rest.

Each repetition was simultaneously captured by two 8-camera markerless motion capture systems (DARI Motion, DARI Motion, Inc., Overland Park, KS, 66210, USA), each collecting at 240 Hz as shown in Figures 1-2. No post-processing synchronization of the two systems was necessary due to the repetition-level nature of the analysis and the native capabilities of the software's algorithms. Variable selection was inclusive of both upper and lower body joints as well as the three cardinal planes of the body for thorough assessment of the algorithm's capabilities. The lower body variables of interest were hip flexion, knee flexion, ankle dorsiflexion, and knee valgus; the upper body variables of interest were shoulder external rotation, shoulder internal

Table 1: List of movements performed

Shoulder Internal/External Rotation
Shoulder Flexion
Trunk Rotation
Overhead Squat
Lateral Lunge
Unilateral Squat
Vertical Jump
Lateral Bound
Unilateral Jump
Unilateral Hop (5 hops)

rotation, and thoracic rotation. Jump height was also considered under the three relevant dynamic movements (i.e., vertical jump, unilateral vertical jump, unilateral five hop). Values of whole body and joint-specific kinematic variables relevant to sports practitioners were processed using the motion capture system's algorithms and exported to a flat file for analysis. Peak values were used for all variables for each subject with the exception of the movement requiring five successive unilateral hops, where the mean jump height was used for subsequent analysis.

3. Statistical Analysis

As explained in Section 2, each subject performed the respective movements for two sets, or repetitions, with each set captured by two systems simultaneously. This resulted in data that could be used to perform four different reliability studies: two inter-rater reliability studies (i.e., the agreement between the systems on each of the sets) and two intra-rater reliability studies (i.e., the agreement between the first and second set for each system). The reliability was assessed for 39 kinematic variables (Table 2) using the ICC and the typical error (TE). Specifically, the form of the appropriate ICC for all four studies is notated as ICC(2, 1) as determined by the flowchart and formulas given in Table 3 of Koo and Li.[10] Furthermore, the TE was computed as the residual standard error in a linear model with subjects and systems (for the inter-rater studies) and sets (for the intra-rater studies) as effects.[7] All data cleaning and analyses were performed using the statistical software environment R.[20] In particular, psych[21] and psr[13] were used to compute the ICCs and TEs, respectively.

Figure 1: The setup for data collection for the current study with 8 cameras for each system positioned approximately at the vertices and edge midpoints of a square. Participants stand in the green-tiled area.

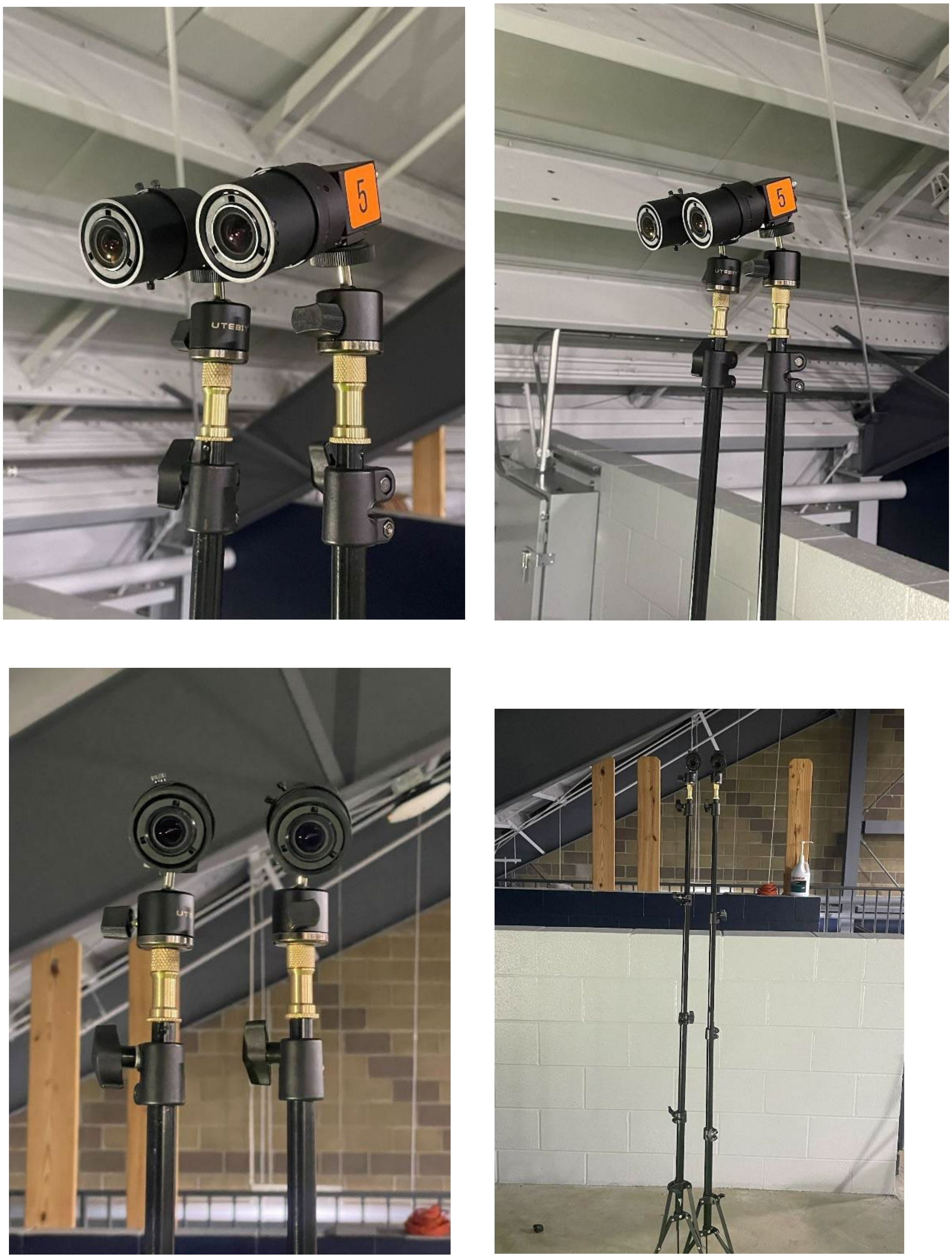

Figure 2: Views of a pair of cameras, one for each system, used for data collection.

Table 2: List of 39 kinematic variables assessed

Shoulder	External Rotation Max: Left & Right
	Internal Rotation Max: Left & Right
	Flexion Max: Left & Right
Thorax	Rotation Max: Left & Right
Hip	Flexion Max Overhead Squat: Left & Right
	Flexion Max Unilateral Squat: Left & Right
	Flexion Max Lateral Lunge: Left & Right
	Hip Abduction Max Overhead Squat: Left & Right
Knee	Flexion Max Overhead Squat: Left & Right
	Flexion Max Unilateral Squat: Left & Right
	Flexion Max Lateral Lunge: Left & Right
	Valgus Angle Max Overhead Squat: Left & Right
	Valus Angle Max Unilateral Squat: Left & Right
	Valgus Angle Max Lateral Lunge: Left & Right
Ankle	Flexion Max Overhead Squat: Left & Right
	Flexion Max Unilateral Squat: Left & Right
	Flexion Max Lateral Lunge: Left & Right
Jump Height	Vertical Jump Height
	Unilateral Jump Height: Left & Right
	Average Hop Height: Left & Right

4. Results

Missing values and outliers were examined first. Each kinematic variable was measured 84 times (21 subjects x 2 systems x 2 sets), and vertical jump height, left unilateral, and right unilateral vertical jump height were missing 12, 21, 30 times, respectively. Additionally, outliers were identified by standardizing each variable and filtering for values greater than 4 or less than -4. A total of eight measurements were flagged as outliers, two each for the variables left and right max knee valgus angle overhead squat and left and right max knee valgus angle unilateral squat. Since outliers were only observed for these variables and the system operators did not report any irregularities in the testing procedure, no observations were removed from the analysis.

The full results for the four reliability studies consisting of 39 kinematic variables are provided in the Appendix (Tables A1-A5). Here, to simplify the presentation, we focus on the results for a subset of the variables. Since many variables had both a "right" and "left" versions (e.g., right and left maximum shoulder flexion), only the results for the right side are presented here. Since the labeling of the systems was arbitrary, we present the results of the first system for the intra-rater study. Finally, since the subjects were more familiar with the protocol after the first set and to account for a potential increase in tissue compliance or treatment effect that would be reflected in an increase in joint range in the subsequent trial, we present the results of the second set for the inter-rater study. Additionally, a pervious study[19] examining reliability of dorsiflexion for a weight bearing lunge include 3-5 trials, which is consistent with being in the second set for this study.

Choosing a threshold value for a minimum acceptable ICC is subjective and may differ depending on the context. In the field of exercise science, Baumgartner and Chung[2] state a minimally acceptable ICC value of 0.70, which has been used in subsequent research into the reliability of sports science technology.[8] We focus first on the inter-rater reliability results. Table 3 shows that variables involving the shoulder were highly reliable (ICC ≥ 0.97) and the CIs for these ICCs were narrow. Examining Table 3 further, the variables concerning thoracic rotation and

Table 3: Inter- and Intra-rater reliability results for shoulder, thorax, and hip

	Inter-Rater			Intra-Rater		
Kinematic Variable	ICC	CI	TE	ICC	CI	TE
Shoulder.External.Rotation.Max	0.98	(0.94, 0.99)	1.66	0.90	(0.75, 0.96)	4.39
Shoulder.Internal.Rotation.Max	1.00	(0.99, 1.00)	1.37	0.59	(0.23, 0.81)	13.69
Shoulder.Flexion.Max	0.97	(0.89, 0.99)	2.03	0.83	(0.64, 0.93)	5.42
Thoracic.Rotation.Max	0.92	(0.68, 0.97)	1.79	0.78	(0.54, 0.90)	3.86
Hip.Flexion.Max.Overhead.Squat	0.89	(0.75, 0.95)	4.21	0.78	(0.54, 0.90)	5.90
Hip.Flexion.Max.Unilateral.Squat	0.91	(0.80, 0.96)	4.80	0.76	(0.51, 0.90)	8.23
Hip.Flexion.Max.Lateral.Lunge	0.94	(0.86, 0.97)	4.73	0.55	(0.18, 0.79)	11.21
Hip.Abduction.Max.Overhead.Squat	0.87	(0.05, 0.97)	1.15	0.80	(0.56, 0.91)	2.04

Table 4: Inter- and Intra-rater reliability results for knee, ankle, and jump height

	Inter-Rater			Intra-Rater		
Kinematic Variable	ICC	CI	TE	ICC	CI	TE
Knee.Flexion.Max.Overhead.Squat	0.99	(0.98, 1.00)	1.12	0.95	(0.87, 0.98)	2.82
Knee.Flexion.Max.Unilateral.Squat	0.99	(0.98, 1.00)	1.18	0.81	(0.46, 0.93)	5.65
Knee.Flexion.Max.Lateral.Lunge	0.99	(0.98, 1.00)	1.17	0.86	(0.61, 0.95)	5.15
Knee.Valgus.Angle.Max.Overhead.Squat	0.79	(0.56, 0.91)	0.12	0.37	(-0.05, 0.68)	0.37
Knee.Valgus.Angle.Max.Unilateral.Squat	0.82	(0.62, 0.92)	0.27	0.85	(0.68, 0.94)	0.22
Knee.Valgus.Angle.Max.Lateral.Lunge	0.59	(0.24, 0.81)	1.21	0.43	(0.02, 0.72)	1.10
Ankle.Flexion.Max.Overhead.Squat	0.90	(0.78, 0.96)	1.69	0.81	(0.59, 0.92)	2.32
Ankle.Flexion.Max.Unilateral.Squat	0.88	(0.72, 0.95)	2.08	0.70	(0.41, 0.87)	3.59
Ankle.Flexion.Max.Lateral.Lunge	0.93	(0.84, 0.97)	2.19	0.68	(0.36, 0.86)	4.29
Jump.Height.Vertical.Jump	1.00	(0.99, 1.00)	0.75	0.97	(0.91, 0.99)	2.06
Jump.Height.Unilateral.Vertical.Jump	1.00	(0.99, 1.00)	0.49	0.96	(0.90, 0.98)	1.92
Average.Jump.Height.5.Hop	0.99	(0.98, 1.00)	0.64	0.91	(0.80, 0.96)	2.15

hip flexion all had an estimated ICC of at least 0.89 and the lower bound of all their respective CIs were above 0.68. Table 4 shows that the variables for knee flexion were highly reliable with CIs (0.98, 1.00). However, the knee valgus variables showed lower reliability, with the estimate for Knee.Valgus.Angle.Max.Lateral.Lunge being 0.59. The middle of Table 4 displays the ICCs for the ankle flexion variables where the ICC estimates are all 0.88 or greater, and the lower bound for all the CIs is greater than 0.70. Finally, the variables for jump height in bottom of Table 4 all show high ICCs and narrow CIs despite the missing data described above.

Turning to the intra-rater reliability results, we see that the ICCs are generally lower than for the inter-rater study. Of the 20 variables in this subset, 14 of them have CI lower bounds below 0.70. The jump height variables showed the closest correspondence between the intra-rater ICCs (i.e., ICC > 0.91) and the corresponding inter-rater ICCs (i.e., ICC > 0.98) than any other group of variables.

5. Discussion

The purpose of the current study was to investigate the inter- and intra-rater reliability of a commercially available markerless motion capture system and the respective algorithms used to derive whole body and joint-specific kinematic variables. The physical setup of the two systems are shown in Figures 1-2. The results of this study show that markerless motion capture is reliable when comparing the results of multiple systems or multiple collection periods. Specifically, regarding inter-rater reliability, the markerless motion capture system provided at least moderate agreement across all variables (i.e., ICC > 0.50) with a vast majority demonstrating excellent reliability (i.e., ICC > 0.90). In particular, inter-rater reliability was excellent across all upper body measurements, inclusive of multi-planar shoulder movements and thoracic rotation. Similarly, inter-rater reliability was largely good for lower body measurements (i.e., ICC > 0.75) except for hip flexion during the overhead squat (ICC = 0.66-0.74) and knee valgus during the overhead squat (ICC = 0.55-0.79), the unilateral squat (ICC = 0.60-0.83), and the lateral lunge (ICC = 0.59-0.89) where the lowest values were interpreted as moderate. However, hip flexion's absolute inter-rater reliability (TE = 3.782-7.657) makes this less concerning for practitioners as normative values for flexion range of motion are likely more than 100 degrees.[24] Knee valgus typical errors were also favorable, with the largest TE observed being 1.205 degrees with respect to inter-rater reliability. In other words, the ICCs for hip flexion or knee valgus were not as strong as other variables, but the absolute agreement between devices makes integration into athlete monitoring processes reasonable.

Intra-rater reliability of the system again ranged from moderate to excellent for all movements and corresponding variables. The majority of variables presented moderate intra-rater reliability, but intra-rater reliability for knee valgus during the overhead squat, unilateral squat, and lateral lunge were mostly poor. However, intra-rater reliability for knee valgus ranged widely, even exceeding ICCs of 0.75 (i.e., good) during the overhead squat and lateral lunge. Generally, intra-rater reliability was lower than inter-rater reliability for the same measurement of consideration, although this could be partly attributed to familiarization or warm-up effect between the first and second sets. Supporting this potential explanation, the intra-rater absolute agreement of shoulder internal rotation was much larger (TE = 11.345-13.685 deg) relative to inter-rater (TE = 1.595-2.056 deg). Furthermore, the respective intra-rater ICC values were similar comparing the two markerless motion capture systems across most variables. Therefore, practitioners using a markerless motion capture system should be familiar with absolute measures of reliability (e.g., typical error) for each variable so that they can better discern whether changes following intervention can be truly attributed to the support provided or normal measurement variability.[15]

It is worth emphasizing that the data should be considered within the context of the joint or nature of the movement. For example, knee valgus angles during movement tasks will be influenced by anatomical properties of the individual (e.g., femur length), making normative values more difficult to establish.[16] Therefore, practitioners intervening in hopes of reducing knee valgus should pursue directionality (i.e., reduction in knee valgus) and change relative to the individual's baseline in lieu of settling into a normative range. With such a conservative typical error (TE = 0.225-1.706 deg), practitioners could reasonably apply the current study's markerless motion capture system within that scope. Alternatively, despite moderate-to-good intra-rater agreement, ankle dorsiflexion demonstrated TE as high as 4.288 deg. Hence, determining meaningfulness of change following intervention in a clinical or sports performance setting will be more difficult considering the similarity between the observed TE and the minimum detectable change at the ankle joint.[15,19]

Both inter- and intra-rater reliability were excellent for both the bilateral and unilateral vertical jumps. Furthermore, the TE ranged from 0.336 to 2.064 cm which is consistent with other common means of approximating center of mass displacement during jumping tasks using flight time[17] and the impulse-momentum relationship.[14] Therefore, the markerless motion capture system is a satisfactory tool for assessing jumping abilities. Notably, the percentage of the number of jump movement repetitions missed by the system during the current study was 14-36% across the bilateral vertical jump and right and left unilateral vertical jumps. Although the absences did not follow a recognizable pattern, they could be explained by performing three jumps in succession without adequate time between repetitions, as the processing algorithm expects a trial with a single jump. Alongside the excellent relative and strong absolute agreement of jump height, it is highly likely that more time between repetitions would reduce or resolve the missingness issue observed in the current investigation. Alternatively, the markerless motion capture algorithm may be challenged during rapid movements. For example, the missingness observed may be the result of identifying when the subjects are disengaged with the floor between takeoff and landing, not permitting approximation of flight times (and therefore jump height). This comment is speculative since the authors are not privy to the proprietary features or methods of calculating jump height. Nonetheless, practitioners are encouraged to collect multiple trials of each jump type with adequate rest between repetitions to account for the missingness observed in the current study. Due to the nature of the data, the potential influence of missing repetitions on the measurement of jump height during five sequential unilateral hops is unknown. Different from all other variables which used peak values, data was reported as an average across a maximum of five trials with no indication to the number of repetitions used (e.g., only three repetitions used in presented averages due to missed repetitions).

6. Conclusions

Markerless motion capture gives sport practitioners the ability to describe movement in terms of kinematics, which is useful in screening or technical development. The results of the current study indicate that markerless motion capture is reliable when comparing the results of multiple systems or multiple collection periods. However, practitioners should consider both relative and absolute agreement as that is critical in applying the information provided by markerless systems. Ultimately, markerless motion capture will progress to capturing the sport movements themselves. For now, it appears to provide a suitable alternative to subjectively scored screening protocols, an advantage brought about by the granularity afforded in a markerless motion capture system. Reliably describing joint position during a movement screen allows for more objective evaluation of the direction and subsequent success of interventions. This may not only have athlete health benefits but may also position coaches to guide technical development with greater precision.

Undoubtedly, markerless motion capture will continue to scale its presence in sport. Future research should explore the potential improvements in reliability and accuracy of these systems, especially when capturing movements with high segment or system velocities.

Disclosure Statement

We declare that we have no conflicts of interest in the authorship or publication of this contribution.

Acknowledgments

The authors would like to thank DARI Motion for providing access to the two systems used in this study.

References

1. Agami, S., Riemer, R., Berman, S. Enhancing motion tracking accuracy of a low-cost 3D video sensor using a biomechanical model, sensor fusion, and deep learning. *Front Rehabil Sci* 2022;3:956381.
2. Baumgartner, T., Chung, H. Confidence limits for intraclass reliability coefficients. *Meas Phys Educ Exerc Sci 2001;5*(3):179-188.
3. Bonett, D. G. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med 2002;21*(9):1331-1335.

4. Federolf, P. A Novel Approach to Solve the "Missing Marker Problem" in Marker-Based Motion Analysis That Exploits the Segment Coordination Patterns in Multi-Limb Motion Data. *PLoS One* 2013;*8*(10):e78689.
5. Fleisig, G., Slowik, J. S., Wassom, D., Yanagita, Y., Bishop, J., Diffendaffer, A. Z. Comparison of Marker and Marker-less Motion Capture for Baseball Pitching Kinematics. *Sports Biomech* 2022;*PMID 35591756.*
6. Hando, B. R., Scott, W. C., Bryant, J. F., Tchandja, J. N., Scott, R. M., Angadi, S. S. Association Between Markerless Motion Capture Screenings and Musculoskeletal Injury Risk for Military Trainees: A Large Cohort and Reliability Study. *Orthop J Sports Med* 2021;*PMID 34734097.*
7. Hopkins, W. G. Measures of reliability in sports medicine and science. *Sports Med 2000;30*(1):1-15.
8. Hori, N., Newton, R., Kawamori, N., McGuigan, M., Kraemer, W. Reliability of performance measurements derived from ground reaction force data during countermovement jump and the influence of sampling frequency. *J Strength Cond Res* 2009;*23*(3):874-882.
9. Kanko, R. M., Laende, E. K., Davis, E. M., Selbie, W. S., Deluzio, K. J. Concurrent assessment of gait kinematics using marker-based and markerless motion capture. *J Biomech* 2021;*127:*110665.
10. Koo, T., Li, M. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-163.
11. Liu G., McMillan, L. Estimation of missing markers in human motion capture. *Vis Comput* 2006;*22*(9-11):721-728.

12. Martinez, H. R., Garcia-Sarreon, A., Camara-Lemarroy, C., Salazar, F., Guerrero-González, M. L. Accuracy of Markerless 3D Motion Capture Evaluation to Differentiate between On/Off Status in Parkinson's Disease after Deep Brain Stimulation. *Parkinsons Dis 2018*;2018:5830364.
13. McGinn, T., Huebner, A., Sisk, M. *psr: Functions for Analyzing Performance Science Data*. R package version 0.1.0. 2021. https://cran.r-project.org/web/packages/psr/index.html.
14. Merrigan, J. J., Stone, J. D., Hornsby, W. G., Hagen, J. A. Identifying Reliable and Relatable Force–Time Metrics in Athletes-Considerations for the Isometric Mid-Thigh Pull and Countermovement Jump. *Sports 2020;9*(1):4
15. Michlovitz, S. L., Harris, B. A., Watkins, M. P. Therapy Interventions for Improving Joint Range of Motion: A Systematic Review. *J Hand Ther* 2004;*17*(2):118-131.
16. Nilstad, A., Krosshaug, T., Mok, K. M., Bahr, R., Andersen, T. E. Association between anatomical characteristics, knee laxity, muscle strength, and peak knee valgus during vertical drop-jump landings. *J Orthop Sports Phys Ther* 2015;45(12):998-1005.
17. Nuzzo, J. L., Anning, J. H., Scharfenberg, J. M. The reliability of three devices used for measuring vertical jump height. *J Strength Cond Res* 2011;25(9):2580-2590.
18. Perrott, M. A., Pizzari, T., Cook, J., McClelland, J. A. Comparison of lower limb and trunk kinematics between markerless and marker-based motion capture systems. *Gait Posture 2017;*52:57-61.
19. Powden, C. J., Hoch, J. M., Hoch, M. C. Reliability and minimal detectable change of the weight-bearing lunge test: A systematic review. *Man Ther* 2015;20(4):524-32.
20. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. URL https://www.R-project.org/.

21. Revelle, W. *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 2.2.9. 2022. https://CRAN.R-project.org/package=psych.

22. Schurr, S. A., Marshall, A. N., Resch, J. E., Saliba, S. A. Two-dimensional video analysis is comparable to 3D motion capture in lower extremity movement assessment. *Int J Sports Phys Ther* 2017;12(2):163-172.

23. Sonnenfeld, J. J., Crutchfield, C. R., Swindell, H. W., Schwarz, W. J., Trofa, D. P, Ahmad, C. S., Lynch, T. S. An Analysis of InVivo Hip Kinematics in Elite Baseball Batters Using a Markerless Motion-Capture System. *Arthrosc Sports Med Rehabil* 2021;3(3):e909-e917.

24. 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;17(3):500-507.

25. Van der Kruk, E., Reijne, M. M. Accuracy of human motion capture systems for sport applications; state-of-the-art review. *Eur J Sport Sci* 2018;18(6):806-819.

26. Whiteside, D., Deneweth, J. M., Pohorence, M. A., Sandoval, B., Russell, J. R., McLean, S. G., Zernicke, R. F., Goulet, G. C. Grading the functional movement screen: A comparison of manual (real-time) and objective methods. *J Strength Cond Res* 2016;30(4):924-933.

Appendix

Table A1: Inter- and Intra-rater reliability results for shoulder

		Inter 1			Inter 2			Intra 1			Intra 2		
Variable		ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE
Ex.Rot.Max	Left	0.96	(0.91, 0.99)	1.98	0.95	(0.88, 0.98)	2.64	0.87	(0.71, 0.95)	3.96	0.88	(0.71, 0.95)	3.55
	Right	0.98	(0.96, 0.99)	1.62	0.98	(0.94, 0.99)	1.66	0.89	(0.75, 0.96)	4.39	0.92	(0.76, 0.97)	3.63
Int. Rot. Max	Left	0.99	(0.98, 1.00)	2.06	0.99	(0.98, 1.00)	1.91	0.73	(0.46, 0.88)	11.35	0.70	(0.40, 0.87)	12.68
	Right	0.99	(0.98, 1.00)	1.60	1.00	(0.99, 1.00)	1.37	0.59	(0.23, 0.81)	13.69	0.53	(0.15, 0.77)	14.50
Flex.Max	Left	0.94	(0.61, 0.98)	2.34	0.94	(0.74, 0.98)	2.89	0.87	(0.70, 0.94)	4.88	0.90	(0.77, 0.96)	4.59
	Right	0.94	(0.80, 0.98)	2.70	0.97	(0.89, 0.99)	2.03	0.83	(0.64, 0.93)	5.42	0.81	(0.59, 0.92)	5.85

Note: Ex.Rot.Max = External.Rotation.Max, Int.Rot.Max = Internal.Rotation.Max, Flex.Max = Flexion.Max

Table A2: Inter- and Intra-rater reliability results for thorax

		Inter 1			Inter 2			Intra 1			Intra 2		
Variable		ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE
Rot.Max	Left	0.95	(0.81, 0.98)	1.88	0.96	(0.85, 0.98)	1.35	0.82	(0.62, 0.92)	3.67	0.71	(0.42, 0.87)	4.31
	Right	0.93	(0.76, 0.98)	1.72	0.92	(0.68, 0.97)	1.79	0.78	(0.54, 0.90)	3.86	0.78	(0.54, 0.90)	3.47

Note: Rot.Max = Rotation.Max

Table A3: Inter- and Intra-rater reliability results for hip

		Inter 1			Inter 2			Intra 1			Intra 2		
Variable		ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE
Fl.Mx.Ov.Sq	Left	0.74	(0.42, 0.89)	6.34	0.83	(0.18, 0.95)	3.78	0.81	(0.60, 0.92)	5.49	0.79	(0.53, 0.91)	5.88
	Right	0.66	(0.34, 0.85)	7.66	0.89	(0.75, 0.95)	4.21	0.78	(0.54, 0.90)	5.90	0.83	(0.61, 0.93)	5.31
Fl.Mx.Uni.Sq	Left	0.97	(0.93, 0.99)	2.81	0.96	(0.74, 0.99)	2.86	0.87	(0.71, 0.95)	6.36	0.85	(0.64, 0.94)	7.13
	Right	0.96	(0.91, 0.98)	3.49	0.91	(0.80, 0.96)	4.80	0.76	(0.51, 0.90)	8.23	0.78	(0.53, 0.90)	8.36
Fl.Mx.Lat.Lu	Left	0.96	(0.88, 0.98)	3.65	0.95	(0.76, 0.98)	3.57	0.89	(0.74, 0.95)	6.50	0.80	(0.57, 0.92)	8.24
	Right	0.92	(0.82, 0.97)	4.60	0.94	(0.86, 0.97)	4.73	0.55	(0.18, 0.79)	11.21	0.71	(0.42, 0.87)	9.42
Ab.Mx.Ov.Sq	Left	0.74	(-0.02, 0.92)	1.60	0.85	(0.11, 0.96)	1.44	0.72	(0.43, 0.87)	2.78	0.78	(0.54, 0.90)	2.52
	Right	0.82	(0.04, 0.95)	1.15	0.87	(0.05, 0.97)	1.25	0.80	(0.56, 0.91)	2.04	0.75	(0.47, 0.89)	2.55

Note: Fl.Mx.Ov.Sq = Flexion.Max.Overhead.Squat, Fl.Mx.Uni.Sq = Flexion.Max.Unilateral.Squat, Fl.Mx.Lat.Lu = Flexion.Max.Lateral.Lunge, Ab.Mx.Ov.Sq = Abduction.Max.Overhead.Squat

Table A4: Inter- and Intra-rater reliability results for knee

		Inter 1			Inter 2			Intra 1			Intra 2		
Variable		ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE
Fl.Mx.Ov.Sq	Left	0.99	(0.97, 1.00)	1.00	0.99	(0.96, 1.00)	1.20	0.91	(0.80, 0.96)	3.68	0.92	(0.82, 0.97)	3.36
	Right	0.99	(0.99, 1.00)	0.88	0.99	(0.98, 1.00)	1.12	0.95	(0.87, 0.98)	2.82	0.94	(0.86, 0.97)	2.96
Fl.Mx.Uni.Sq	Left	1.00	(0.97, 1.00)	0.82	0.99	(0.97, 1.00)	1.28	0.73	(0.42, 0.88)	7.10	0.71	(0.42, 0.87)	7.22
	Right	1.00	(0.99, 1.00)	0.88	0.99	(0.98, 1.00)	1.18	0.81	(0.46, 0.93)	5.65	0.80	(0.43, 0.92)	5.66
Fl.Mx.Lat.Lu	Left	0.99	(0.96, 1.00)	0.89	0.99	(0.91, 1.00)	0.87	0.92	(0.77, 0.97)	4.30	0.91	(0.75, 0.97)	4.39
	Right	0.99	(0.99, 1.00)	1.11	0.99	(0.98, 1.00)	1.17	0.86	(0.61, 0.95)	5.15	0.89	(0.67, 0.96)	4.42
VA.Mx.Ov.Sq	Left	0.94	(0.85, 0.97)	0.38	0.55	(0.18, 0.79)	1.02	0.85	(0.66, 0.93)	0.68	0.55	(0.17, 0.79)	0.85
	Right	0.69	(0.38, 0.86)	0.33	0.79	(0.56, 0.91)	0.12	0.37	(0.00, 0.68	0.37	0.32	(0.00, 0.64)	0.38
VA.Mx.Uni.Sq	Left	0.60	(0.25, 0.81)	0.15	0.60	(0.25, 0.82)	0.58	0.21	(0.00, 0.58)	0.46	0.12	(0.00, 0.51)	0.73
	Right	0.63	(0.28, 0.83)	0.31	0.82	(0.62, 0.92)	0.27	0.85	(0.68, 0.94)	0.22	0.51	(0.13, 0.76)	0.42
VA.Mx.Lat.Lu	Left	0.84	(0.65, 0.93)	0.66	0.86	(0.69, 0.94)	1.01	0.49	(0.09, 0.75)	1.71	0.29	(0.00, 0.63)	1.72
	Right	0.89	(0.75, 0.95)	0.46	0.59	(0.24, 0.81)	1.21	0.43	(0.02, 0.72)	1.10	0.39	(0.00, 0.69)	1.32

Note: Fl.Mx.Ov.Sq = Flexion.Max.Overhead.Squat, Fl.Mx.Uni.Sq = Flexion.Max.Unilateral.Squat, Fl.Mx.Lat.Lu = Flexion.Max.Lateral.Lunge, VA.Mx.Ov.Sq = Valgus.Angle.Max.Overhead.Squat, VA.Mx.Uni.Sq = Valgus.Angle.Max.Unilateral.Squat, VA.Mx.Lat.Lu = Valgus.Angle.Max.Lateral.Lunge

Table A5: Inter- and Intra-rater reliability results for ankle

		Inter 1			Inter 2			Intra 1			Intra 2		
Variable		ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE
Fl.Mx.Ov.Sq	Left	0.90	(0.15, 0.98)	1.23	0.95	(0.88, 0.98)	1.55	0.89	(0.75, 0.96)	2.00	0.89	(0.75, 0.95)	2.19
	Right	0.86	(0.69, 0.94)	1.89	0.90	(0.78, 0.96)	1.69	0.81	(0.59, 0.92)	2.32	0.76	(0.50, 0.89)	2.50
Fl.Mx.Uni.Sq	Left	0.92	(0.81, 0.97)	1.90	0.93	(0.83, 0.97)	1.58	0.80	(0.58, 0.91)	2.85	0.85	(0.66, 0.93)	2.55
	Right	0.89	(0.74, 0.95)	2.20	0.88	(0.72, 0.95)	2.08	0.70	(0.41, 0.87)	3.59	0.78	(0.54, 0.90)	2.83
Fl.Mx.Lat.Lu	Left	0.94	(0.86, 0.97)	2.06	0.97	(0.92, 0.99)	1.38	0.78	(0.50, 0.91)	3.28	0.65	(0.32, 0.84)	4.44
	Right	0.93	(0.83, 0.97)	1.94	0.93	(0.84, 0.97)	2.19	0.68	(0.36, 0.86)	4.29	0.76	(0.40, 0.90)	3.27

Note: Fl.Mx.Ov.Sq = Flexion.Max.Overhead.Squat, Fl.Mx.Uni.Sq = Flexion.Max.Unilateral.Squat, Fl.Mx.Lat.Lu = Flexion.Max.Lateral.Lunge

Table A6: Inter- and Intra-rater reliability results for jump height

		Inter 1			Inter 2			Intra 1			Intra 2		
Variable		ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE	ICC	ICC CI	TE
Vertical		1.00	(1.00, 1.00)	0.44	1.00	(0.99, 1.00)	0.75	0.97	(0.91, 0.99)	2.06	0.96	(0.91, 0.99)	2.36
Uni.Vert	Left	1.00	(0.99, 1.00)	0.55	0.99	(0.99, 1.00)	0.58	0.99	(0.98, 1.00)	0.71	0.98	(0.90, 1.00)	0.87
	Right	1.00	(1.00, 1.00)	0.34	1.00	(0.99, 1.00)	0.49	0.96	(0.90, 0.98)	1.92	0.96	(0.91, 0.99)	1.64
Avg.5.Hop	Left	0.97	(0.93, 0.99)	1.26	0.99	(0.98, 1.00)	0.62	0.85	(0.67, 0.94)	2.89	0.85	(0.64, 0.94)	2.82
	Right	1.00	(0.99, 1.00)	0.42	0.99	(0.98, 1.00)	0.64	0.91	(0.80, 0.96)	2.15	0.90	(0.78, 0.96)	2.32

Note: Vertical = Vertical jump, Uni.Vert = Unilateral vertical jump, Avg.5.Hop = Average over 5 hops