AGREEMENT BETWEEN A SMARTPHONE-BASED MARKERLESS MOTION CAPTURE APPLICATION (AISCOUT ®) AND CONVENTIONAL ATHLETIC ASSESSMENT

Simon Augustus[1], James Brouner[1], Chris Howe[2], Anna Kosciuk[2], & Richard Felton-Thomas[2]

[1] Department of Applied & Human Sciences, Kingston University London, UK
[2] Project 23rd Century Ltd. (ai.io), London, UK

There is a need to develop accessible 3D motion analysis. The aim of this study was to determine absolute and relative agreement between aiScout®, a smartphone application that infers 3D pose from 2D video, and conventional methods of assessing athletic performance. Sixteen athletes (29.7 ± 9.0 yrs, 1.75 ± 0.10 m, 75.5 ± 15.8 kg) completed six performance tests that were evaluated by aiScout® and conventional methods. 10m sprint times, press-ups, lateral rebound jumps, and 5-10-5 football dribble times showed good to excellent absolute (ICC and CCC > 0.8 threshold) and relative agreement (ICC and Kendall's W > 0.8 threshold) between measurement methods. Unilateral countermovement jumps only showed moderate agreement (0.6 - 0.8 absolute and relative agreement). This suggests aiScout® can be used to help assess athletic performance. However, questions remain regarding the capability of aiScout® to assess performance of jumping activities.

KEYWORDS: artificial intelligence, 3D motion analysis, accessibility

INTRODUCTION: Traditional three-dimensional motion analysis (3DMA) uses marker-based systems to reconstruct position and orientation of human movement. These provide valuable insights regarding execution of sports skills, but require expensive multi-camera set-ups, intrusive marker attachments and considerable expertise for operation and data analysis. Inertial measurement units have also been used for 3DMA, but similar problems remain. High financial costs and lack of expertise ultimately preclude athletes of lower socio-economic status from benefiting from 3DMA. This indicates a need to develop cost-effective and accessible solutions. Given a) global proliferation of smart devices and high-quality digital video technology and b) advances in computer vision and artificial intelligence (AI), precise and reliable 3DMA from a single camera view should be an important and attainable goal.
An alternative is markerless motion analysis. These systems remove long preparation times and marker attachments, are commercially (e.g. Theia Inc) and freely available (e.g. OpenCap), and have been validated against traditional marker-based systems for activities such as walking gait (Kanko et al., 2021) and squatting (Ulrich et al., 2023). Whilst readily performed with multiple cameras, reconstructing 3D human pose from a single camera is more challenging. Computer vision and AI techniques have proved promising and recent techniques have been introduced as proof of concept (e.g. Pavlakos et al., 2018). Unfortunately, for applied scientists and practitioners with limited computer science expertise, these are not easily implemented.
Recently, aiScout®, a computer vision and AI application that offers coaches and athletes the opportunity to perform 3DMA in athletic settings using a smart device (i.e. phone or tablet) was released. The athlete is identified in each video frame and the images converted to 2D pose. Next, 3D pose is inferred using AI algorithms, and kinematic performance metrics extracted. If aiScout® is to be a useful tool for coaches and athletes to help assess athletic performance 'in the field', a necessary first step was to determine if performance metrics derived from aiScout® agree with conventionally used field measurement methods during various athletic tasks (e.g. vertical jumping = force plate, sprinting = timing gates). The aim of this study was thus to determine absolute (i.e. to assess how closely test scores agree) and relative agreement (i.e. to assess how closely do participant rankings agree) between aiScout® and conventional measurement methods for athletic assessment.

METHODS: Sixteen recreational team sport athletes (males = 12, females = 4; mean ± SD; age = 29.7 ± 9.0 years, stature = 1.75 ± 0.10 m, mass 75.5 ± 15.8 kg) were recruited. All provided written informed consent and approval was granted by the University's Ethics Committee. Following warm-up and familiarisation, participants performed six tests chosen to represent a variety of technical and physical athletic tasks performed in different planes and body orientations (one repetition of a 10 m sprint, unilateral countermovement jump (CMJ) left and right leg, 30 second lateral rebound and press-ups, and a 5-10-5 football dribble). These were concurrently assessed using aiScout® and conventional methods. For the 10 m sprint and 5-10-5 dribble tests, time to completion was assessed conventionally using wireless timing gates (Brower SpeedTrap 2; Draper, USA). Unilateral vertical CMJ jump height was assessed by a piezoelectric force platform (1000 Hz; 9286B, Kistler, UK) by integrating CoM acceleration (net vertical force divided by body mass) and using the take-off velocity method. The number of repetitions for 30 second press-up and lateral rebounds were determined using two smartphone videos (sagittal and frontal view; 4K, 60Hz; iPhone 13, USA). For press-ups, repetitions were discounted if: a) elbow angle was > 30° and < 25° at end of the eccentric phase, b) the upper arm did not dip below < 15° relative to horizontal (sagittal plane) during the eccentric phase, c) hip flexion angle was < 150° or > 210°, or d) varied more than 35° throughout the repetition, and e) knee angle was < 150°. For lateral rebounds, participants jumped side-to-side over a football. Repetitions were discounted if the participant did not clear the vertical height or lateral edge of the football. Videos were exported to Kinovea (v0.9.5; Kinovea.org) for manual count of repetitions.

For aiScout® assessment, trials were recorded using a separate smartphone (4K, 60Hz; iPhone 13, USA). Scaling was performed using one or two FIFA approved size 5 footballs (diameter 0.23m) and 'in app' instructions followed to replicate how users would perform tests 'in the field'. Videos were uploaded to aiScout®'s online platform where 3DAT (3D athlete tracking) software extracts an 18-point 2D pose from key body landmarks. Next, a camera-distance aware, top-down, neural network infers 3D pose by estimating the depth of each landmark and scaling objects relative to the camera's lens (Moon et al., 2009). 10 m sprint times, 5-10-5 dribble times and CMJ heights were determined from whole-body centre of mass (CoM) position. Press-up and lateral rebound repetitions were determined using the constraints noted above and by vertical and horizontal line boundaries placed on the left/right and top of ball relative to ankle position. Intraclass correlation coefficients (ICC; two-way mixed effects, absolute agreement, multiple raters; Koo & Li, 2016) and Lin's concordance coefficients (CCC; Lin, 1989) were calculated to assess absolute agreement of test scores. Since these analyses do not account for both fixed and proportional bias, ordinary least products regression (OLPR) was also conducted (Ludbrook, 2012). If the 95% CI for the intercept did not include 0, then fixed bias was present. If the 95% CI for the slope did not include 1, then proportional bias was present. ICC (two-way mixed effects, consistency, multiple raters) and Kendall's *W* concordance coefficient with correction for tied ranks (Legendre, 2010) were calculated for ordinal participant rankings for each test to assess relative agreement. Correlation point estimates and 95% confidence intervals were interpreted as < 0.5 = poor, 0.5 – 0.75 = moderate, 0.75 – 0.9 = good and > 0.9 = excellent agreement (Koo & Li, 2016), and minimum agreement thresholds as 0.8 (Baumgartner & Chung, 2001). All ICCs were conducted in JASP (0.16.3, https://jasp-stats.org/), and CCCs, OLPR and Kendall's W were conducted in Matlab (R2022a; MathWorks, Natick, MS), respectively.

RESULTS: ICCs for absolute agreement and CCCs showed good to excellent agreement for the 10 m sprint, 5-10-5 dribble, press-ups and lateral rebounds, and moderate to excellent agreement for the unilateral CMJs (Table 2). All ICC for absolute agreement point estimates exceeded the minimum agreement threshold, but CCC point estimates and both ICC and CCC lower 95% CIs for the unilateral CMJs did not (Table 2). OLPR indicated there was no fixed or proportional bias (Table 2). For relative agreement, ICCs for consistency and Kendall's *W* showed good to excellent agreement for the 10m sprint, 5-10-5 dribble, press-ups and lateral rebounds (Table 3). The unilateral CMJs showed moderate to excellent agreement, but ICC for consistency point estimates and their lower 95% CIs fell below the agreement threshold.

Table 1. Mean ± SD descriptive data and mean differences (95% CI) for each test.

	Conventional	aiScout®	Mean Difference
10 m Sprint Time (s)	2.09 ± 0.14	2.10 ± 0.13	-0.01 (-0.03, 0.00)
CMJ Height (Left) (m)	0.12 ± 0.04	0.14 ± 0.03	-0.02 (-0.03, -0.01)
CMJ Height (Right Leg) (m)	0.12 ± 0.03	0.14 ± 0.04	-0.02 (-0.03, -0.01)
30 Second Press-Ups Reps	19.3 ± 13.0	19.0 ± 13.0	0.3 (-1.3, 1.9)
30 Second Lateral Rebound Reps	44.8 ± 13.1	45.6 ± 13.7	-0.8 (-1.9, 0.3)
5-10-5 Dribble Times (s)	8.36 ± 1.61	8.47 ± 1.63	-0.20 (-0.30, 0.00)

Table 2. ICC (two-way mixed effects, absolute agreement, multiple raters), CCC and OLPR (95% CI) for assessment of absolute agreement. * = point estimate < 0.8 agreement threshold.

	ICC	CCC	OLPR Intercept	OLPR Slope
10 m Sprint Time (s)	0.981 (0.946, 0.993)	0.975 (0.933, 0.990)	0.107 (-0.019, 0.323)	0.956 (0.852, 1.059)
CMJ Height (Left) (m)	0.841 (0.603, 0.941)	0.728* (0.447, 0.879)	2.780 (-0.518, 6.079)	0.917 (0.635, 1.199)
CMJ Jump Height (Right) (m)	0.847 (0.617, 0.944)	0.759* (0.493, 0.900)	-0.179 (-4.376, 4.019)	1.147 (0.806, 1.487)
30 Second Press-Ups Reps	0.975 (0.930, 0.991)	0.973 (0.832, 0.958)	-0.679 (-3.254, 1.897)	1.019 (0.886, 1.152)
30 Second Lateral Rebound Reps	0.988 (0.967, 0.996)	0.987 (0.964, 0.995)	-1.378 (-5.261, 2.507	1.049 (0.962, 1.136)
5-10-5 Dribble Time (s)	0.992 (0.978, 0.997)	0.990 (0.973, 0.997)	-0.001 (-0.596, 0.594)	1.013 (0.635, 1.084)

Table 3. ICC for consistency (two-way mixed effects, consistency, multiple raters) and Kendall's *W* (95% CI) for assessment of relative agreement. * = point estimate < 0.8 agreement threshold.

	ICC	Kendall's *W*
10 m Sprint Time Rank	0.962 (0.894, 0.986)	0.977 (0.933, 0.992)
CMJ Height (Left) Rank	0.690* (0.311, 0.879)	0.833 (0.575, 0.940)
CMJ Height (Right) Rank	0.599* (0.164, 0.839)	0.812 (0.530, 932)
30 Second Press-Ups Rank	0.939 (0.834, 0.978)	0.973 (0.922,0.991)
30 Second Lateral Rebound Rank	0.974 (0.928, 0.991)	0.989 (0.968, 0.996)
5-10-5 Dribble Rank	0.989 (0.969, 0.996)	0.994 (0.982, 0.998)

DISCUSSION: The 10 m sprint, 5-10-5 football dribble, press-up and lateral rebound tests all showed good to excellent absolute agreement between measurement methods, and there was no evidence of fixed or proportional bias. For these tests, aiScout® provided performance metrics that were comparable to those derived from conventional assessment. Good to excellent agreement was also shown in terms of participant rankings (relative agreement), suggesting aiScout® can reliably distinguished between high and low performers in these tests. The results add to a growing body of research that has used AI to conduct valid field-based evaluation of sports skills (Webering et al., 2021) and offers preliminary indication that 3D kinematics derived from a single camera view can be successfully utilised in athletic contexts. Previous studies have offered similar solutions using computer vision and AI, but these have not been evaluated empirically (e.g. Pavlakos et al., 2018). The present results concur with research that showed excellent agreement between the 'SpeedClock' smartphone application and timing gates for assessing sprint performance (Stanton et al., 2016). Unfortunately, it is not clear what methods SpeedClock uses to derive sprint times. Since no study has evaluated football dribbling, press-up and lateral rebound performance using AI, evaluation of these data are difficult. One observation was that aiScout® underestimated press-up repetitions for one participant (difference in reps = 9). Further inspection showed this participant performed press-ups 'out of plane' to the camera used to derive aiScout® metrics. More shoulder abduction/ adduction rather than flexion/ extension meant the upper arm and shoulder became obscured by the elbow during the eccentric phase. aiScout® was thus unable to distinguish if the upper arm had dipped below the threshold of < 15° relative to horizontal. In contrast, conventional assessment utilised the second frontal camera to identify this angle. Future research might investigate limitations of aiScout® for reliably identifying kinematics during non-planar activities that obscure skeletal key points from view. Unilateral CMJs showed only moderate to excellent absolute agreement in jump heights between aiScout® and the force platform method, with some ICC and CCC values and their lower 95% CIs falling below the agreement threshold. Caution is thus needed when evaluating aiScout®'s assessment of unilateral CMJ performance. Given the relatively small variation (Table 1) and that ICCs for relative agreement showed the lowest levels of agreement, it was difficult to reliably distinguish between high and low performers. Discrepancies in jump heights derived from kinematic (aiScout®) and kinetic (force platform) methods are well known. The former determines jump height from an estimate of CoM position, whereas the latter uses integration of the vertical force-time curve. Kinematic methods suffer from error associated with determination of segment orientations, whereas force platforms suffer from 'drift' introduced by the integration process (Wade et al., 2020). These issues are worse in unilateral CMJs, as it is difficult to obtain an adequate period of quite standing. It is important to be aware of the shortcomings of both aiScout® and force platform methods of determining jump height. Nevertheless, the level of agreement is comparable to studies that have performed similar analyses. Future research is warranted to further evaluate the validity of aiScout® for determining CMJ kinematics. Comparison of aiScout® against a marker-based 3DMA system might offer opportunities to improve the system. This study chose to focus on derivation of key performance metrics that will be used to directly inform athletic assessment 'in the field'. A necessary next step would be to evaluate the precision of joint and segment kinematics derived from aiScout® across a wider range of athletic activities. This will enable more detailed assessment of aiScout's® capability to provide valid and reliable 3DMA.

CONCLUSIONS: For most tests, aiScout® showed good to excellent absolute and relative agreement, and an absence of fixed and proportional bias when compared to conventional assessment of athletic performance. Preliminary evidence is thus provided for the utility of aiScout® as a useful tool for coaches and athletes to help assess athletic performance. However, questions remain regarding the capability of aiScout® to assess jumping activities. To move towards development of cost-effective and accessible 3DMA solutions using a single 2D camera, future research should evaluate if 3D kinematics derived from aiScout® can provide precise assessment of joint and segment level kinematics against gold-standard motion analysis techniques (e.g. discrete and time-series joint angles and velocities).

REFERENCES

Baumgartner T.A., & Chung H. (2001). Confidence Limits for Intraclass Reliability Coefficients. Measurement in Physical Education and Exercise Science, (5) 179–88.

Kanko, R., Laende, E., Davis, E., et al. (2021). Concurrent assessment of gait kinematics using marker-based and markerless motion capture. Journal of Biomechanics, 127, 110665.

Koo, T.K, &, Li M.Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine. 15, 155–163.

Legendre, P. (2010). Coefficient of Concordance. In: Salind NJ (Ed), Encyclopedia of research design. SAGE.

Lin, L. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics. 45, 255–268.

Ludbrook J (2012) A primer for biomedical scientists on how to execute Model II linear regression analysis. Clinical and Experimental Pharmacology and Physiology, (39)329–35.

Pavlakos, G., Zhu, L., Zhou, X., et al. (2018). Learning to Estimate 3D Human Pose and Shape from a Single Color Image (arXiv:1805.04092).

Stanton, R., Hayman, M., Humphris, N., et al. (2016). Validity of a Smartphone-Based Application for Determining Sprinting Performance. *Journal of Sports Medicine*, 1–5.

Uhlrich, S. D., Falisse, A., Kidziński, Ł., et al. (2023). OpenCap: Human movement dynamics from smartphone videos. *PLOS Computational Biology*, *19*, e1011462.

Wade, L., Lichtwark, G., & Farris, D. (2020). Comparisons of laboratory-based methods to calculate jump height and improvements to the field-based flight-time method. *Scandinavian Journal of Medicine & Science in Sports*, *30*, 31–37.

Webering, F., Blume, H., & Allaham, I. (2021). Markerless camera-based vertical jump height measurement using OpenPose. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 3863–3869.