Validation of PITCHAI™ Markerless Motion Capture Using Gold Standard 3D Motion Capture

Supplementary materials:
www.osf.io/xxx
For correspondence:
mholmes2@brocku.ca

Tyler J Dobos[1], Ryan WG Bench[1], Michael WR Holmes[1]* Colin D McKinnon[2], Anthony Brady[3], Kyle J Boddy[3], Michael WL Sonne[1,2]

[1]Department of Kinesiology, Faculty of Applied Health Sciences, Brock University, St. Catharines, ON, Canada

[2]3MotionAI, Oakville, ON, Canada

[3]Driveline Baseball LLC, Kent, WA, USA

Please cite as: Dobos et al. (2022). Validation of PITCHAI™ Markerless Motion Capture Using Gold Standard 3D Motion Capture. *SportRχiv*.

ABSTRACT

Most kinematic and kinetic assessments in baseball pitchers have been determined using marker-based motion capture systems. No current research exists on the feasibility of using single camera markerless motion capture technology for kinematic analysis of pitching. The purpose of this study was to compare and validate pitching kinematics (joint angles and summary metrics) from a markerless motion capture solution with a gold standard, 3D optical marker-based solution. 38 healthy pitchers of high school, college, or professional levels of experience threw 1-3 maximum effort pitches while concurrently using marker-based optical motion capture and pitchAI smartphone based (markerless) motion capture. Each pitch was time normalised from peak leg lift, to ball release. Time series pitchAI measures were compared to 3D motion capture using Pearson's R (R), R

All authors have read and approved this version of the manuscript. This article was last modified on January 8, 2022.

Authors Dobos @drtylerdobos can be reached on Twitter.

Squared (r^2), and root mean square error (RMSE) for each joint angle. Discrete time points were evaluated for all joint kinematics at foot plant (FP), maximal shoulder external rotation (MER), and ball release (BR); as well as for descriptive metrics (stride length, arm speed, ball visibility, and ball path). For full time-series joint angles the pelvis and trunk had the best overall fit with an average r^2 of 0.98, and 3.1 ± 1.1° of RMSE. The knee angles had an average r^2 of 0.97 ± 0.02, and an average RMSE of 4.1 ± 2.0°. The throwing arm had an average r^2 of 0.97 ± 0.02, with an average RMSE of 6.0 ± 2.2° across all measures. The glove arm performed the worst, with an average r^2 of 0.95, and 7.3° of RMSE. Across all discrete time points, the most accurate measures were the knees, followed by the trunk and pelvis, throwing arm, and finally glove arm. Stride length had an average RMSE of 3.90 ± 4.77%, and an r^2 of 0.31. Arm speed had an average RMSE of 2.6 ± 3.4 m/s, and an r^2 of 0.25. Ball visibility had an RMSE of 20.7 ± 24.1 ms, and an r^2 of 0.10. Ball path had an RMSE of 19.79 ± 23.9%, and an r^2 of 0.45. When considering the technical ease of video-based solutions and an ability to measure in field, most metrics were within an acceptable range to the gold standard. pitchAI can be recommended as a markerless alternative to classic marker-based motion capture for baseball pitch kinematic analysis.

Key Words: kinematics; baseball; pitching; motion capture

INTRODUCTION

Technological advancements within the motion capture field have enabled athletes and coaches to accurately quantify and easily capture the movement demands of their sport (Mündermann et al., 2006; Boddy et al., 2019). The most common methods for accurate capture and assessment of three-dimensional human movement require a laboratory environment and the attachment of markers, fixtures or sensors to body segments; where infrared cameras track reflective markers in 3D space with sub-millimeter accuracy (Fleisig et al., 1996; Barrentine et al., 1998; Richards, 1999; Mündermann et al., 2006; Thewlis et al., 2013). In the sport of baseball, professional organizations, research institutes, and private training facilities have been making use of motion capture techniques since the early 1990's. Baseball pitchers tend to demonstrate a prototypical delivery to throw the baseball which can be divided into phases allowing a better understanding of the biomechanics of the body as the pitch is delivered. The pitching phases were originally described as the windup, stride, early arm cocking, late arm cocking, arm acceleration, arm deceleration, and the follow through (Dillman et al., 1993; Werner et al., 1993). In general, foot plant (FP) identifies the end of the stride and early arm cocking phases and beginning of the late cocking and arm acceleration phases. Next within the arm acceleration phase, is the late cocking phase, which ends at maximum external rotation (MER). Arm acceleration ends at ball release (BR), which is then followed by arm deceleration and follow through (Dillman et al., 1993; Werner et al., 1993). With foot plant, maximum external rotation, and ball release

being markers of the beginning and end of phases, they are common time points to investigate and report postures and angular velocities of athletes. Although, using discrete time points does not give insight into how an athlete is moving from position to position, time series data provides important information that may indicate more efficient deliveries (Fleisig et al., 1999). If a single camera motion capture solution could provide valid and reliable data with minimal hardware, this would allow coaches and researchers to obtain biomechanics data in a wide array of settings for insight into fatigue, player development and performance.

Biomechanical assessments of sport performance often require technical knowledge, setup experience, expensive equipment, can be time-intensive, and require controlled conditions to maximize reliability (Nicholls et al., 2003; Kanko et al., 2021). While standards vary, a typical marker-based motion capture solution for tracking human movement during baseball pitching can require greater than 48 markers being placed on the body (Fleisig et al., 1996; Nicholls et al., 2003). While this can be time intensive and require expertise, controlled laboratory environments provide a high level of accuracy (Fleisig et al., 1996; Barrentine et al., 1998; Mündermann et al., 2006). However, camera placements, consistent marker placement, marker movement artifacts, and clothing restrictions make marker-based motion capture difficult in situ (i.e., in game, on field) (Fleisig et al., 1996; Nicholls et al., 2003). Due to the unique environments posed during data collection, pitchers may not throw "naturally", or how they might in a game situation. As a result, investigations of pitching kinematics in a lab environment have pitchers throwing into a net or a target, and not a live-game setting (Fleisig et al., 1996; Fleisig et al., 2009; Escamilla et al., 1998; Stodden et al., 2005). Such factors restrict the number of assessments that can be performed at a time, and can reduce access and extrapolation of the information. Additionally, these technological limitations can make it difficult for scientists to recruit elite athletic populations. With the exception of a few elite training and research facilities, many studies report on amateur athletic populations. Further, most biomechanics assessments represent a snapshot in time with only a few maximal intensity pitches recorded (Fleisig et al., 1995, Fleisig et al., 1996, Fleisig et al., 1999, Fleisig et al., 2009; Stodden et al., 2005). This limits longitudinal inference on pitching performance (such as kinematic changes to maintain velocity during fatigue (Birfer et al., 2019)) and injury progression. Kinematic changes occur as a result of fatigue (Grantham et al., 2014; Yang et al., 2014), but our understanding is limited to pre-post investigations with little insight into the progression of changes. Improved access would enable pitching coach's greater insight for individualized coaching and feedback; and would shine light onto the understudied effects of fatigue or performance related mechanical changes. Despite the aforementioned drawbacks that accompany gold standard marker-based motion capture, it is still the most accurate and reliable method for tracking 3D human movement to date.

Advances in video-based (markerless) motion capture for pose estimation (Desmarais et al., 2021; Kanko et al., 2021; Vafadar et al., 2021) could offer solutions for

many of the challenges identified with marker-based solutions. Being able to capture and analyze kinematics using a video camera can simplify the technical aspects of motion capture. Further, increased access to video-based kinematic analysis could lead to more research on higher skilled populations, in-game situations, and the tracking of fatigue during competition. Lately, automatic human pose estimation using deep learning strategies have garnered interest amongst computer vision researchers (Ionescu et al., 2014; Chen et al., 2017). The majority of these algorithms prime a neural network using manually labeled image databases (ie., Human3.6M, ENSAM) and then estimate human posture, including anatomical joint centers, when the user inputs the images or videos to the trained network (Nakano et al., 2020; Vafadar et al., 2021). Such software has previously been demonstrated as an effective approach for measuring joint angles in three dimensions during simulated occupational tasks (McKinnon et al., 2019).

The purpose of this study was to compare and validate pitching kinematics (joint angle and summary metric) from a markerless motion capture solution with a gold standard, 3D optical marker-based solution. The markerless motion capture solution used here is pitchAI (3MotionAI Inc, Oakville, ON, Canada). pitchAI estimates 3D joint angles from sagittal plane video of a pitcher, and trains its machine learning models from gold standard marker-based data, to track positional data of anatomical landmarks during a baseball pitch. pitchAI calculates joint angles and angular velocities during a baseball pitch, identifying phases of the pitch in a standardized biomechanics report based on the pitching motion. We hypothesize joint angle measurements from pitchAI would be accurate and reliable compared to gold standard motion capture. More specifically, better agreement would be found for measurements such as elbow flexion, shoulder abduction, shoulder horizontal abduction, and knee flexion than for more complex coordinate system measurements such as pelvis rotation, trunk measurements, and shoulder external rotation.

METHOD

Participants and Ethical Approval

38 healthy pitchers of high school, college, or professional level experience from Driveline Baseball participated in the study. Data collection occurred at a Driveline Baseball Arizona-based motion capture facility. Approval for this study and data collection was provided by the Brock University Bioscience Research Ethics Board (REB #19-371). Prior to data collection, the study was verbally explained to all subjects, and an informed consent was explained and signed. Subject participation was part of their normal training environment during preparation for the 2020 summer season. Demographic information including height, weight, handedness, and level of experience was recorded. Players were included in this study if they were capable of pitching maximum effort for 1-3 pitches, pitched from an overhead armslot, and played competitive baseball. Players were excluded if they pitched “sidearm” or a ¾ armslot, or were incapable of throwing maximum effort for all captured

pitches due to pain or injury. Individual pitches were excluded if they were not able to be processed by pitchAI for reasons such as incomplete capture (recording started early or late).

Experimental Set-up

Pitchers were permitted to perform a self-directed warm up so they were ready to perform between 1-3 maximum effort throws. Pitch velocities were recorded using a Stalker Pro 2 doppler radar gun (Applied Concepts/Stalker Radar, Richardson, TX, USA). Throws were made using a five-oz. (142 g) regulation baseball from a regulation size mound to a strike zone target (Oates Specialties, LLC, Huntsville, TX, USA) located above home plate, which was 60′ 6″ (18.4 m) away. Pitches were recorded concurrently using marker-based optical motion capture and video camera. Video was time-synced with the motion capture using the peak lead leg knee marker in the vertical direction. The coordinate system for both markered and markerless motion capture was aligned and a laboratory coordinate system was established based on International Society of Biomechanics standards: Y was vertical, X was perpendicular to Y (positive to home plate), and Z was orthogonal to Y and X (Wu et al., 2002; Wu et al., 2005).

Markered Motion Capture

Pitchers were fitted with reflective markers in preparation for their throwing data collection. A total of 48 reflective markers were attached bilaterally on the third distal phalanx, lateral and medial malleolus, calcaneus, tibia, lateral and medial femoral epicondyle, femur, anterior and posterior iliac spine, iliac crest, acromial joint, inferior angle of the scapula, midpoint of the humerus, lateral and medial humeral epicondyle, midpoint of the ulna, radial styloid, ulnar styloid, distal end of index metacarpal, parietal bone, and frontal bone, as well as on the C7 and T10 vertebrae, the sternal end of the clavicle, and the xiphoid process (Brady et al., 2020). Markers were tracked at 240Hz using an 11 camera OptiTrack system (NaturalPoint Inc, OptiTrack, Corvallis, OR, USA). This system contained a mixture of 7, Prime 13 and 4, Prime 13W cameras. Cameras were placed symmetrically around the capture volume on tripods, approximately 8–12 feet from the center of the pitching mound at varying heights. The tracked outputs recorded using the OptiTrack system were processed in Visual3D (C-Motion Inc., Germantown, MD, USA) using a fourth order zero-lag Butterworth digital low-pass filter with a cut-off frequency of 20Hz. All joint centre data were exported for further kinematic analysis (Figure 1).

Markerless Motion Capture

Video was captured using an iPhone 11 (Apple Inc, Cupertino, CA, USA) at 240Hz, 1080p resolution, with the iPhone handheld by one of the investigators. Video was recorded in the native camera app from an open-side angle in landscape mode, with the pitcher held in frame throughout the entirety of the throwing motion. The video was uploaded to pitchAI's

source code for offline processing. Pose estimation was performed via pitchAI, with step 1 being a generation of a 2D, 19 joint center model. Position data were filtered using a fourth order zero-lag Butterworth digital low-pass filter with a cut-off frequency of 13.4 Hz (Werner et al., 1993), and gaps interpolated using an Akima spline fill. A neural network (38 input nodes; a single 100-node hidden layer - rectified linear unit activation; 159 output nodes - linear activation) was developed as a corrective measure for the markerless motion capture data based on the joint center position data from the marker-based solution. The pitchAI neural network transformed the 2D, 19 marker, 38 point output data into a 3D, 19 joint center, 53 marker, 159 xyz point output. Estimated marker locations included 19 joint centers: head, neck, and pelvis, plus bilateral shoulder, elbow, wrist, hip, knee, ankle, heel, and toe markers; in addition to 34 bony landmarks: C7, upper sternum, T8, and the xiphoid process as well as bilateral anterior and posterior iliac spine, lateral head, acromion, medial epicondyle, lateral epicondyle, ulnar styloid, radial styloid, hand, medial femoral condyle, lateral femoral condyle, lateral malleolus, medial malleolus, heel, and toe markers. Motion capture marker trajectories were used to construct forearm, upper arm, thigh, shank, foot, pelvis, thorax, and head segment local coordinate systems.

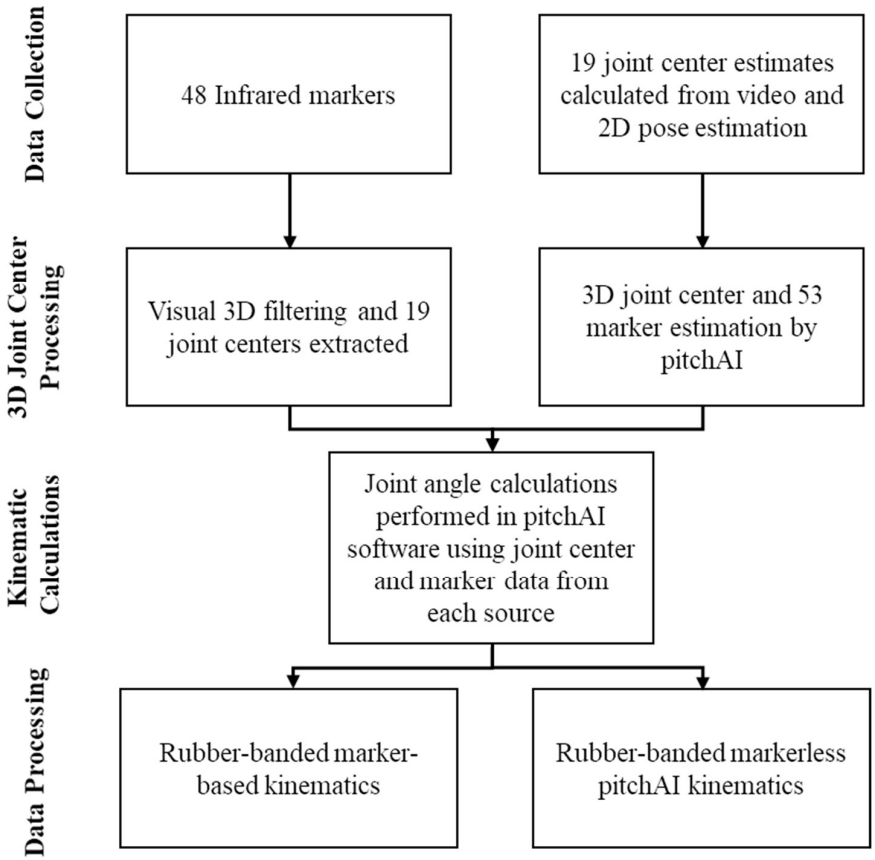

Figure 1. Data collection and processing steps for preparing markered and markerless data for kinematic analysis. Irrespective of data source, joint angle calculations were performed using the same method.

Kinematic analysis

Time-series joint angle data were calculated from both markered and markereless data collection methods, using established methods to approximate joint centers based on the location of the markers on bony landmarks mentioned above. The joint centers were approximated as follows, with the shoulder joint center offset from acromial joint marker (Veeger, 2000), elbow joint center between medial and lateral humeral epicondyle markers, wrist joint center between radial and ulnar styloid markers, hip joint based on pelvis markers, knee joint between medial and lateral femoral condyle markers, ankle joint between medial and lateral malleolus markers, toe joint center at the toe marker, and heel at the heel marker. Local coordinate systems were used to generate all segment and joint angles based on ISB standards for 3-axis Euler angle calculation.

Data analysis

Each pitch was time normalised from the start (peak vertical knee position) to the end of the pitch (ball release, plus an additional 10% of that time after ball release). Time series pitchAI measures were compared against the motion capture tracked measures using Pearson's R (R), R-Squared (r^2), and root mean square error (RMSE) separately for each calculated joint angle. Angles included: 1) throwing arm elbow flexion (defined as the angle between vectors created from the shoulder to elbow, and elbow to wrist; 0° was full elbow extension), 2) throwing arm shoulder horizontal abduction (angle between vectors created from shoulder to shoulder, and shoulder to elbow; 0° was in the frontal plane), 3) throwing arm shoulder abduction (angle between vectors created from hip to shoulder, and shoulder to elbow; 0° was arm at the side of the torso), 4) throwing arm shoulder external rotation (the angle between planes created using the shoulder, elbow and wrist, relative to the plane of the torso using both shoulders, and hips; 0° was at 90° shoulder abduction, 90° elbow flexion with the hand facing forward), 5) glove arm elbow flexion (resolved the same as the throwing arm), 6) glove arm shoulder horizontal abduction (resolved the same as the throwing arm), 7) glove arm shoulder abduction (resolved the same as the throwing arm), 8) glove arm shoulder external rotation (resolved the same as the throwing arm), 9) rear leg knee extension (angle between vectors created from hip to knee, and knee to ankle; 0° was full knee extension), 10) lead leg knee extension (resolved the same as the rear leg), 11) trunk forward tilt (angle between vector from center pelvis, to center upper torso, relative to vertical, resolved in the sagittal plane; 0° was vertical), 12) trunk lateral tilt (angle between vector from centre pelvis, to center upper torso, relative to vertical, resolved in frontal plane; 0° was vertical), 13) trunk twist (angle between vector from shoulder to shoulder, relative to the plane from the mound to home plate; 0° was shoulders in line with home plate), 14) pelvic twist (angle between vector from hip to hip, relative to the plane from the mound to home plate; 0° was hips in line with home plate), and 15) hip-shoulder separation (the difference between the trunk and pelvis twist angles; 0° was when the vectors were aligned). Descriptive metrics included: 1) stride length (as a percentage of

height, measured from ankle joint center to ankle joint center), 2) arm speed (peak throwing arm wrist velocity in the direction of home plate), 3) ball visibility (the time that a ball would be visible to a batter between foot plant and ball release), and 4) ball path (the distance the ball travels from start of pitch to foot plant). RMSE of the two time series data were used as the indicator of the difference as described by Equation (1), where, n is the number of frames, y and yi are the positions estimated by the marker-based and markerless approaches, respectively.

$$RMS\ Error = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}i - yi)^2}{n}} \qquad (1)$$

The r, r^2, and RMSE was also calculated for the kinematic variables at the time points of foot plant (FP), maximum external rotation (MER), and ball release (BR). FP was defined as the peak deceleration of the lead ankle in the y (vertical) direction. MER was defined as the peak throwing arm shoulder external rotation angle between foot plant and ball release. BR was defined as the first zero crossing after the peak acceleration of the wrist in the forward direction, plus an additional 0.035 seconds after this time. All statistical analyses, including r^2, and RMSE, were performed in Microsoft Excel, version 16.52 (Microsoft, Redmond, WA, USA).

Results

Evaluation of the time series, average joint angles obtained by the two different motion capture methods are presented in Table 1. Figure 2 demonstrates example time series comparisons of shoulder external rotation and trunk rotation for both motion capture methods. The average discrete time-point joint angles are depicted in Figure 3 for FP, Figure 4 for MER and Figure 5 for BR. Descriptive metrics are depicted in Figure 6.

Full time-series joint angles

The pelvis and trunk had the best overall fit, with an average r^2 of 0.98, and 3.1° of RMSE throughout the entire signal. The glove arm performed the worst, with an average r^2 of 0.95 and 7.3° of RMSE. Individually, the best performing metrics were the trunk and pelvis twist, with 3.2° and 2.9° error, respectively. The worst performing metric was the glove arm external rotation, with 10.5° of error. All values can be found in Table 1.

Table 1. Time series analysis for joint angles between the marker-based motion capture system (Optitrack), and video-based pitchAI. Statistics are correlation values, as well as RMS error, and % error.

	Throwing Arm				Glove Arm				Lead Leg	Trail Leg	Pelvis and Trunk				
	Elbow Flexion	Shoulder Horizontal Abduction	Shoulder Abduction	Shoulder External Rotation	Elbow Flexion	Shoulder Horizontal Abduction	Shoulder Abduction	Shoulder External Rotation	Knee Extension	Knee Extension	Forward Flexion	Lateral Tilt	Trunk Twist	Pelvis Twist	Hip-Shoulder Separation
r	0.98	0.98	1.00	0.98	0.98	0.98	0.99	0.94	0.99	0.98	0.99	0.98	1.00	1.00	0.99
r^2	0.96	0.96	0.99	0.96	0.95	0.96	0.99	0.89	0.99	0.96	0.98	0.96	1.00	1.00	0.98
RMS error (°)	6.61	7.31	2.76	7.24	5.98	10.26	2.4	10.46	5.53	2.75	2.40	2.22	3.15	2.86	4.93

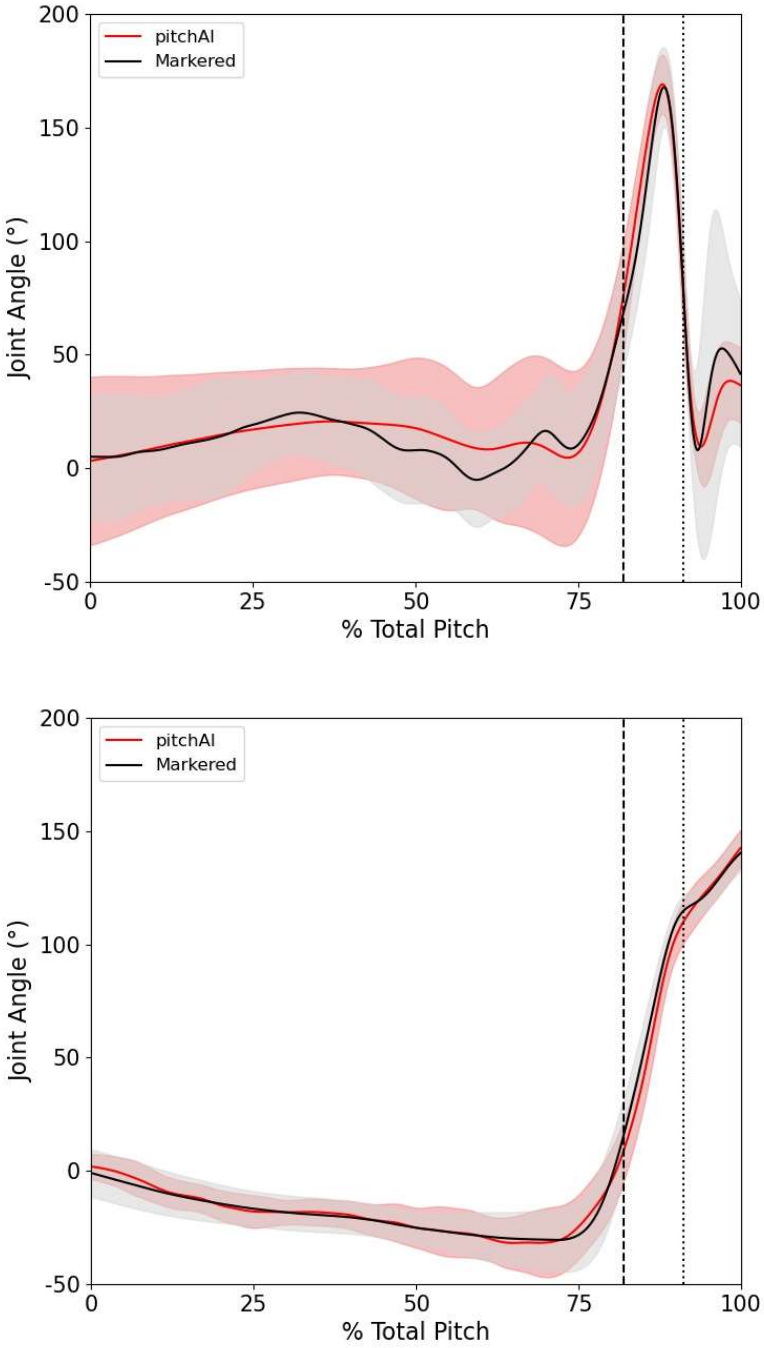

Figure 2. Representative data for two calculated metrics. Time series joint angles (degrees) for shoulder external rotation (top) and trunk rotation (bottom), reported by pitchAI (red), and marker-based motion capture (black). Shaded area represents standard deviation and the data is time normalized as a percentage of the total pitch. Dashed vertical lines indicate FP and BR.

Foot Plant (FP)

At the instant of FP during the pitch, knee angle metrics were the most accurate, with an average RMSE of 7.1 ± 6.5° and an average r^2 of 0.60 ± 0.40. Measures of the trunk and pelvis had an average RMSE of 8.4 ± 4.8°, with an average r^2 of 0.60 ± 0.40. The throwing arm had an average RMSE of 13.0 ± 5.5°, with an average r^2 of 0.58 ± 0.31. Finally, the glove arm had an average RMSE 17.5 ± 6.5°, an average r^2 0.22 ± 0.38. Descriptive statistics and individual metric values can be seen in Figure 3.

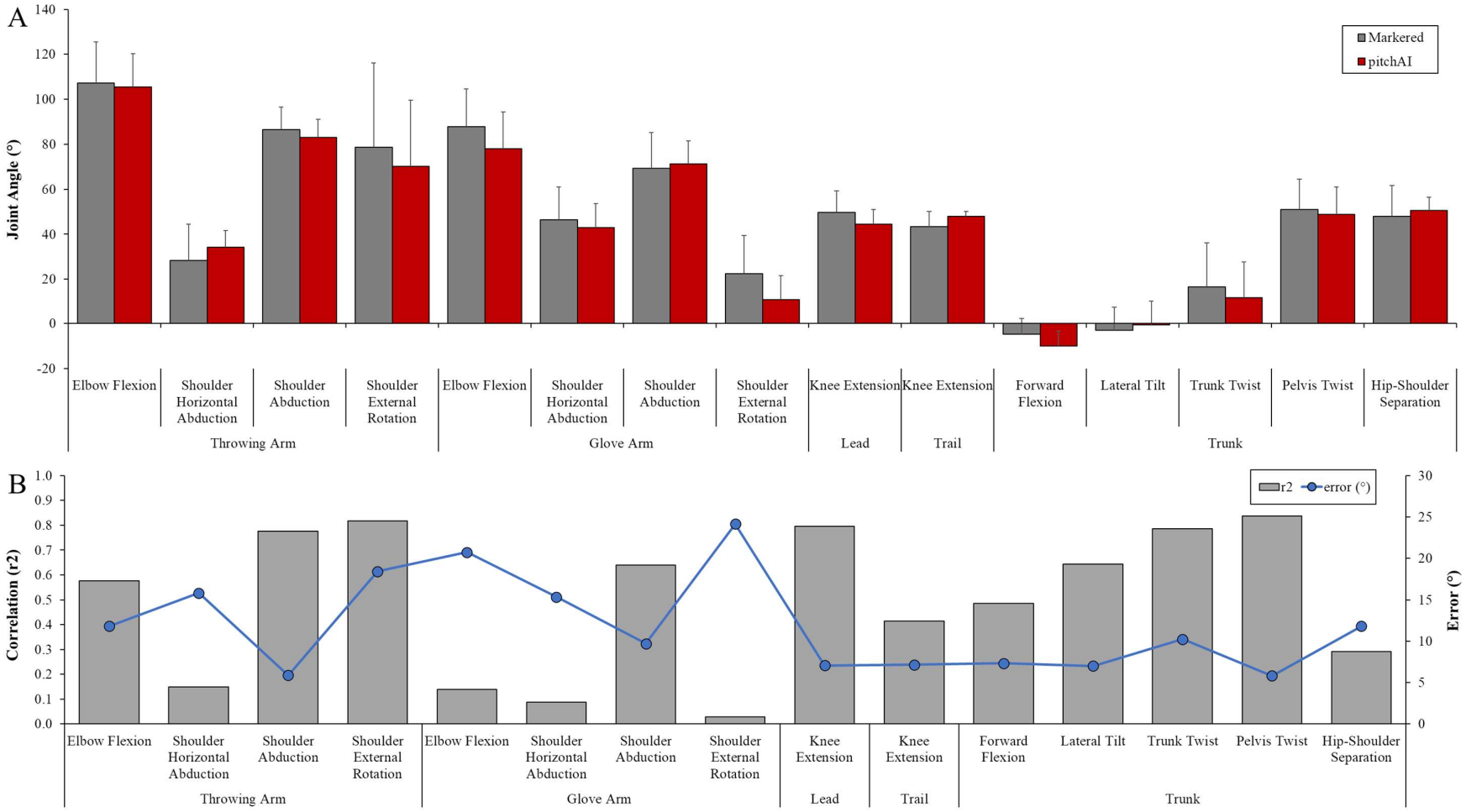

Figure 3. A) Average (SD) joint angles reported by pitchAI and marker-based motion capture at FP. B) Correlation (r^2) and RMS error (°) at FP.

Maximum External Rotation (MER)

At the instant of MER during the pitch, knee angle metrics had an average RMSE of 9.2 ± 3.2°, with an average r^2 of 0.89 ± 0.19. The trunk and pelvis had an average RMSE of 9.6 ± 4.9°, an average r^2 of 0.18 ± 0.22. The throwing arm had an average RMSE of 14.0 ± 2.7°, an average and an average r^2 of 0.37 ± 0.10. The glove arm had an average RMSE of 14.5 ± 2.1° and an average r^2 of 0.24 ± 0.10. Descriptive statistics and individual metric values can be seen in Figure 4.

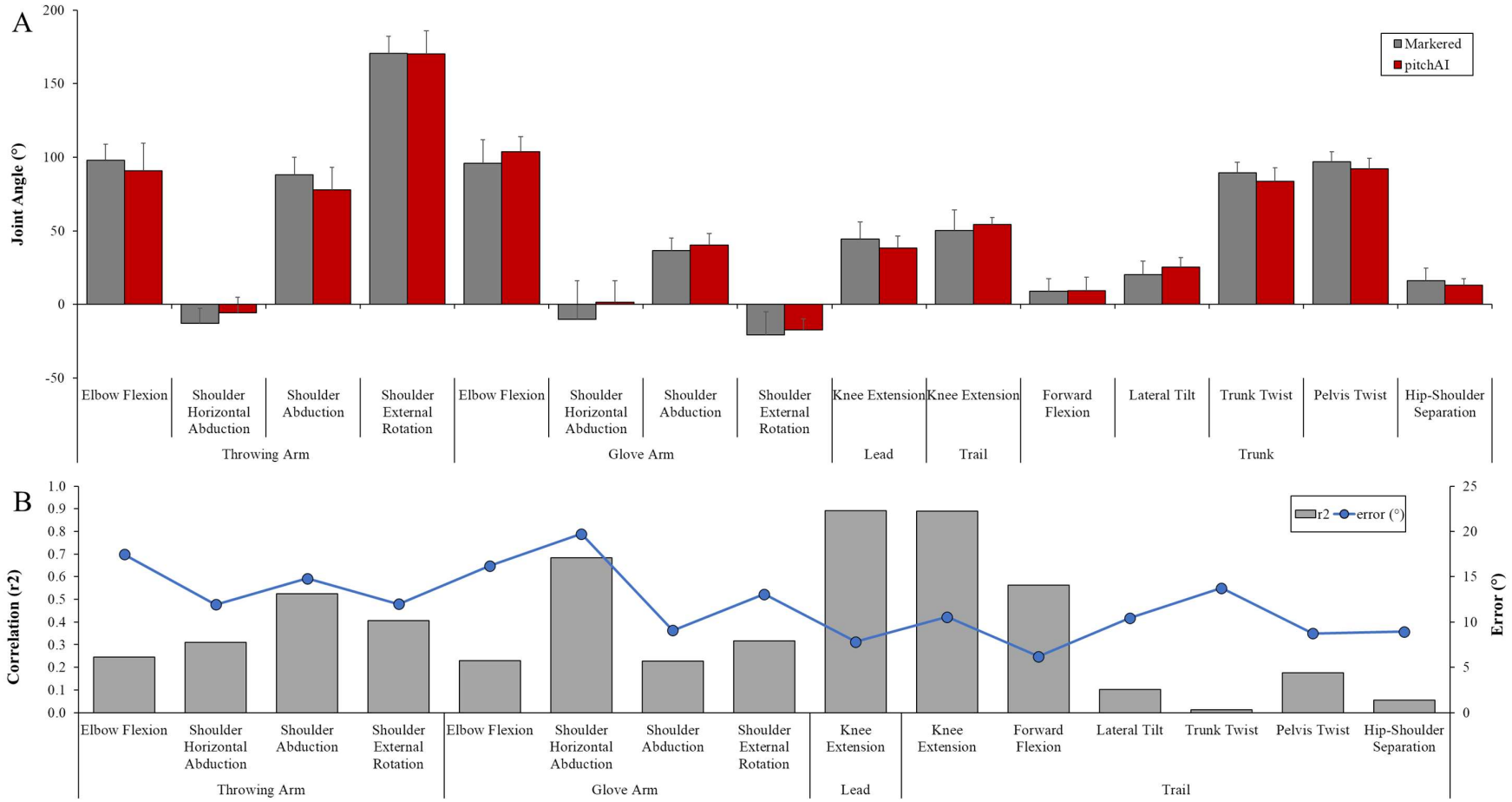

Figure 4. A) Average (SD) joint angles reported by pitchAI and marker-based motion capture at MER. B) Correlation (r^2) and RMS error (°) at MER.

Ball Release (BR)

At the instant of BR during the pitch, the knee angle metrics had an average RMSE of 11.1 ± 3.3°, and an average r^2 of 0.77 ± 0.21. The trunk and pelvis had an average RMSE of 9.9 ± 6.3°, and an average r^2 of 0.22 ± 0.21. The throwing arm had an average RMSE of 15.2 ± 4.0°, and an average r^2 of 0.33 ± 0.16. The glove arm had an average RMSE of 17.2 ± 3.9°, with an average r2 of 0.24 ± 0.22. Descriptive statistics and individual metric values can be seen in Figure 5.

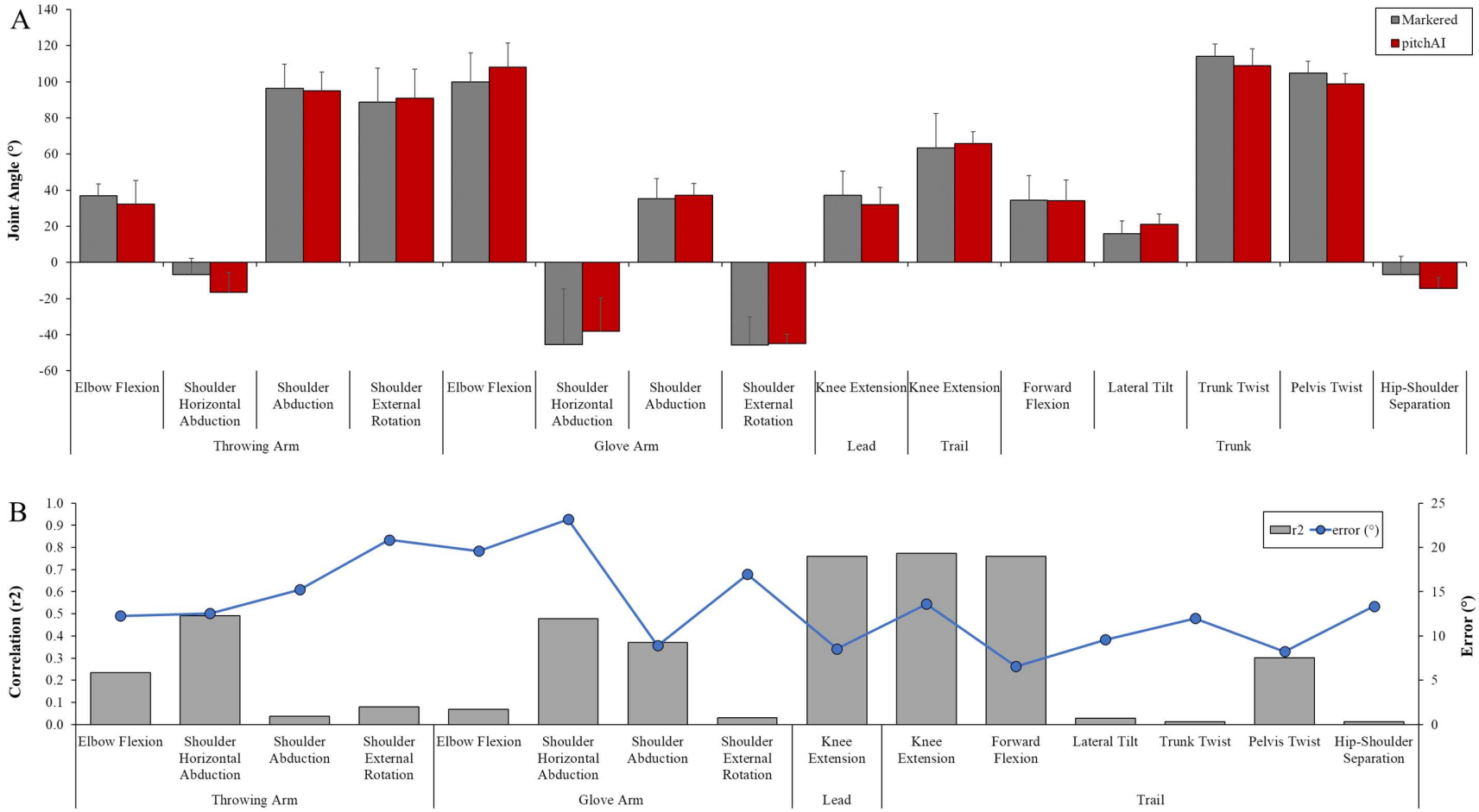

Figure 5. A) Average (SD) joint angles reported by pitchAI and marker-based motion capture at BR. B) Correlation (r^2) and RMS Error (°) at BR.

Descriptive Metrics

Stride length had an average RMSE of 3.9 ± 4.8%, an error of 24%, and an r^2 of 0.31. Arm speed had an average RMSE of 2.6 ± 3.4 m/s, and an r^2 of 0.25. The ball visible time metric had an RMSE of 20.7 ± 24.1 ms, and r^2 of 0.10. The ball path metric had an RMSE of 19.8 ± 23.9%, and an r^2 of 0.45. Descriptive statistics and metrics can be seen in Figure 6.

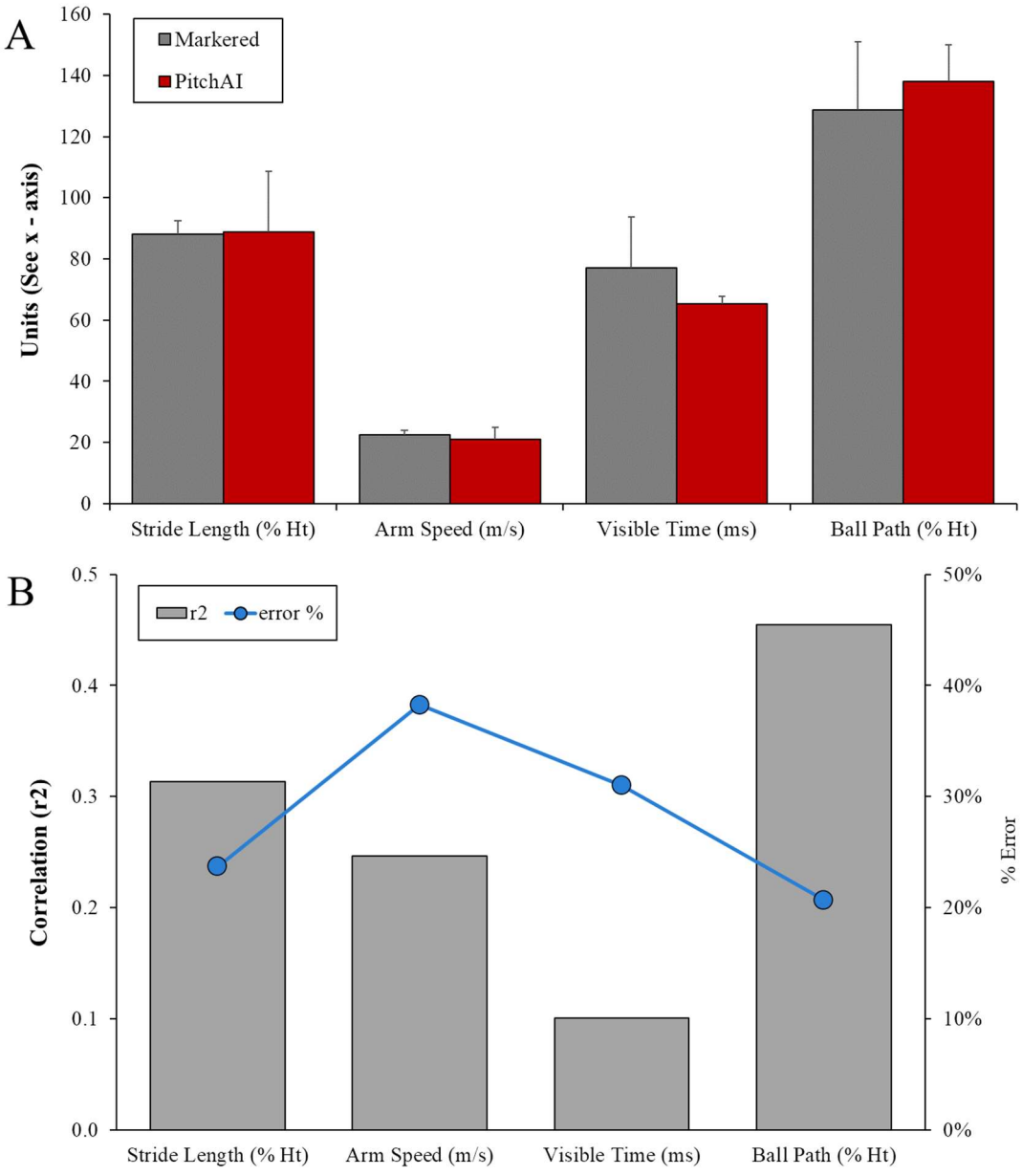

Figure 6. A) Average (SD) descriptive metrics reported by pitchAI, and marker-based motion capture. B) Correlation (r^2) and % error. % Error is calculated as the RMS error divided by the maximum - minimum value for that metric.

Discussion

The purpose of this study was to compare joint angle measures and descriptive metrics from a computer vision-based pose estimation system (pitchAI), against a marker-based motion capture system across a range of joints during the baseball pitching motion. In this study, both motion capture coordinate data were processed using pitchAI, to allow for consistent comparison between sources. For both the full signal analysis, as well as the sequenced markers of FP, MER, and BR, consistent error patterns existed. r^2 values for all joint angles except glove arm shoulder external rotation were greater than 95%, with throwing arm shoulder abduction angle, glove arm shoulder abduction angle, and lead knee extension angle calculated at 99%. Trunk twist angle and pelvis twist angle were 100%. Glove arm shoulder external rotation had an r^2 of 89%, making it the weakest joint angle estimate of the entire regression model. These values demonstrate the strength of the model against gold standard motion capture data. The hypothesis of this work, that better agreement would be found for vector measurements compared to complex coordinate system measurements, was partially supported; with lower agreement for glove arm shoulder external rotation angle and higher agreement for vector measures, but strong agreement for throwing arm shoulder external rotation angle and perfect agreement found for complex measures of trunk twist angle and pelvis twist angle. Overall, the digital video-based measures created from pitchAI were a good fit with those obtained by gold standard motion capture. pitchAI can be recommended as a markerless alternative to classic marker-based motion capture for baseball pitch kinematic analysis, with cautious exception to glove arm shoulder external rotation.

In a motion capture laboratory setting, multiple cameras enable the system to capture 3D coordinates across time and space. The standard capture position for a pitchAI recording is a smartphone in landscape orientation, open-side (predominantly sagittal plane) view. As this is a 2D capture, when the pitcher begins to rotate the pelvis and torso, the glove arm is obstructed from camera view, which decreases the accuracy of joint center estimation. Pose estimation and joint center tracking will be compromised with occlusions or when segments are perpendicular to the capture, as the process would be sensitive to the calculation of image depth when transforming from 2D to 3D space (McKinnon et al., 2020). This was also seen with the throwing arm, as metrics of elbow flexion, shoulder external rotation, and shoulder abduction all had high correlation and low error, but shoulder horizontal abduction performed more poorly as it was the furthest from the plane of assessment. This is a major limitation in single camera, 2D-based transformation engines, such as the one used in this study. However,

when considering these limitations of 2D digital signal interpretation and the issues that arise with the glove arm in general, our reported r^2 of 89% is quite good.

When assessing RMSE of the time-point data across all measures, there was a trend of negative agreement as the throwing motion progressed from peak knee height in the windup, towards ball release (RMSE was 5.8 ± 2.8° at FP; 10.9 ± 4.1° at MER; and 12.3 ± 5.8° at BR). This increase in error can be partly explained through the increasing movement speed that occurs throughout the baseball throwing motion, as potential energy from the beginning of the windup is transferred into kinetic energy that eventually results in the throwing of the baseball (Seroyer et al., 2010). The traditional pitching motion starts slow and gets progressively faster, to ensure that maximum velocity is transferred to the ball. Greater whole-body movement and segmental rotations will result in less data between important phases of the pitch, highlighting the importance of a high sampling rate. This increase in error can also be partly attributed to movement consistency. As a pitcher progresses forward through the delivery, movement variability increases (Scarborough et al., 2020). Many pitchers assume more similar positions early in the throw during the early phases, but share less in common through foot plant and the deceleration phase (Scarborough et al., 2020).

For the current study, video data were captured using a hand-held smartphone, to mimic the typical in field use. The videos were filmed in the native camera app with the subject constantly held in frame, and then processed through pitchAI's source code offline. While the pitchAI application allows for pre-collected slow-motion video to be uploaded and analyzed, it is not the recommended capture method as it is prone to less in-app instruction and greater user error. However, the technology still enables a large "margin of error" within which captures are still possible. Future investigations should determine the effect of camera angle, frame rates, lighting, and pitcher as a percentage of the screen on pitchAI outputs. Nakano et al. (2020) noted that the positions of the landmarks that are tracked with markerless technologies did not correspond identically to the points estimated by a marker-based approach. Similar joint center estimates are used in pitchAI, and therefore to evaluate the accuracy of markerless motion capture, we compared the corresponding positions of the shoulder, elbow, wrist, hip, knee, and ankle joints.

Despite pitchAI providing clear instructions for video film setup, there could be differences that can occur with regard to film angle. This could lead to parallax-induced challenges on system agreement that may produce slightly different results if filmed from a different angle (Tian et al., 2002). One such solution to this problem includes markerless motion capture setups that utilizes multiple camera video captures, such as Theia3D (Theia

Markerless Inc., Kingston, ON), which have previously been shown to be comparable to classical markered solutions (Kanko et al., 2021). Desmarais et al., (2021) reported that when looking at results of contemporary methods, there is an average gap of 10mm error between monocular and multi-view methods. The current process for estimating 3D joint locations from 2D video is still relatively new (McKinnon et al., 2020). Future versions of computer vision-based motion capture will need to consider non perpendicular views to parallel the accuracy and capacity of lab-based systems (McKinnon et al., 2020). It is expected that such 2D-to-3D estimates will undergo rapid enhancements in accuracy over the next decade, which suggests an optimistic outlook for the adaptation of automated, video-based biomechanical analysis (McKinnon et al., 2020).

In addition, our relatively small sample size can explain some of the discrepancies noted between the markerless motion capture and the gold-standard. As the technology becomes more prominent and its user base continues to grow, far greater amounts of data will be available for interpretation and adjustment. A larger sample size would include various different throwing mechanics and deliveries, that may lend themselves to some of the differences seen in this study. The markered motion capture data consisted primarily of overhand or higher three-quarter armslots, with no sidearm or submarine deliveries included in analysis. As a larger and more robust dataset is used to generate a stronger neural network, we expect these values to continue to improve.

Previous reports have described how kinematic changes occur as a result of fatigue (Grantham et al., 2014; Yang et al., 2014; Birfer et al., 2019), however they offer little insight into the progression of changes throughout a game or across a full season. pitchAI would enable pitching coaches and analysts to take intermittent video throughout pitching appearances to track and compare mechanics and metrics from inning to inning, or game to game. This would facilitate active in-season monitoring of player fatigue and performance, which could lend itself to strategic decision making when considering player health and team success.

Perhaps most importantly to note, is that this type of markerless motion capture analysis proposes a more accurate and objective method of performing biomechanical analyses compared to subjective approaches administered by visual inspection. Nicholls et al., (2003) evaluated 17 metrics of the pitching delivery as being “proper or improper”, as evaluated by coaches in a video analysis. These results were compared to actual outputs from a 3D, marker-based motion capture assessment. In terms of agreement, only four of the variables were able to be analyzed properly by the subjective evaluation - elbow flexion at foot contact, sequence of hip-shoulder rotation from the arm cocking phase, and trunk flexion and

horizontal abduction at release. Birfer (2019) evaluated 16 different mechanical properties, and had coaches and scouts subjectively bin body positions into 3 to 5 angle ranges based on previously conducted research. In total, the subjective analysis returned the correct answer on body position 50% of the time - and in many cases, joint angle bins were up to 120 degrees broad. On the spectrum of accessibility and accuracy, this AI-aided video-based approach appears to provide greater accuracy than subjective analysis, while minimizing the amount of technical overhead associated with traditional biomechanical assessments. Multi-camera markered (Richards, 1999; Thewlis et al., 2013) and markerless (Tanaka et al., 2018; Nakano et al., 2020) approaches represent more accurate assessment methods than pitchAI, but when considering the scope of athletes who may be able to afford or access the service, this could be a viable compromise. Advances in computer-vision technology make this type of approach an exciting stop-gap between full lab analysis and show potential for identifying biomechanical issues that could be further analyzed when more technology is available to the pitcher.

Conclusion

The emergence of statistical data surrounding pitch characteristics and pitch velocity have provided deep insight to the baseball community in recent years. However, the biomechanical data behind baseball pitches has thus far remained restricted to a laboratory setting. The results from this study show that a video-based pose estimation approach (i.e. pitchAI) can perform as a low-cost alternative to a fully marker-based kinematic motion capture system, particularly for tracking of pelvis, trunk, and throwing arm metrics in baseball pitching.

Contributions

Contributed to conception and design: All authors
Contributed to acquisition of data: RB, AB, KB, MS
Contributed to analysis and interpretation of data: All authors
Drafted and/or revised the article: All authors
Approved the submitted version for publication: All authors

Acknowledgements

Thank you to the athletes who participated in this study.

Funding information

MH was supported by an NSERC Discovery Grant (RGPIN 2015-05765) and the Canada Research Chairs program (231151 NSERC CRC).

Data and Supplementary Material Accessibility

The raw data supporting the conclusions of this article will be made available by the authors, upon request.

REFERENCES

Barrentine, S. W., Matsuo, T., Escamilla, R. F., Fleisig, G. S., & Andrews, J. R. (1998). Kinematic analysis of the wrist and forearm during baseball pitching. *Journal of Applied Biomechanics*, *14*(1), 24–39.

Birfer, R., Sonne, M. W., & Holmes, M. W. (2019). Manifestations of muscle fatigue in baseball pitchers: A systematic review. *PeerJ*, *7*, e7390.

Birfer, R. (2019). "The Development of a Novel Pitching Assessment Tool. Brock University. MSc Thesis.

Boddy, K., Marsh, J., Caravan, A., Lindley, K., Scheffey, J., & O'Connell, M. (2019). Exploring wearable sensors as an alternative to marker-based motion capture in the pitching delivery. *PeerJ*, *7*. https://doi.org/10.7717/peerj.6365

Brady, A., Briend, S., Caravan, A., Lindley, K., O'Connell, M. E., Gowdey, G., & Boddy, K. J. (2020, February 25). A Kinematic and Kinetic Comparison of Mound and Rocker Throws. https://doi.org/10.31236/osf.io/j2tvg

Cao, Z., Hidalgo Martinez, G., Simon, T., Wei, S. -E., & Sheikh, Y. A. (2019). OpenPose: RealTime Multi-Person 2D pose estimation using part affinity fields. IEEE transactions on pattern analysis and machine intelligence, 15, 1. https://doi.org/10.1109/TPAMI. 2019.2929257

Chen, C. H., and Ramanan, D. (2017). "3D human pose estimation= 2D pose estimation + matching," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (Honolulu, HI), 7035–7043. doi: 10.1109/CVPR.2017.610

Dillman, C. J., Fleisig, G. S., & Andrews, J. R. (1993). Biomechanics of Pitching With Emphasis Upon Shoulder Kinematics. *Journal of Orthopaedic & Sports Physical Therapy*, *18*(2), 402–408. https://doi.org/10.2519/jospt.1993.18.2.402

Desmarais, Y., Mottet, D., Slangen, P., & Montesinos, P. (2021). A review of 3D human pose estimation algorithms for markerless motion capture. *Computer Vision and Image Understanding*, 103275.

Escamilla, R., Fleisig, G., Barrentine, S., Zheng, N., & Andrews, J. (1998). Kinematic Comparisons of Throwing Different Types of Baseball Pitches. *Journal of Applied Biomechanics*, *14*, 1–23. https://doi.org/10.1123/jab.14.1.1

Fleisig, G. S., Andrews, J. R., Dillman, C. J., & Escamilla, R. F. (1995). Kinetics of baseball pitching with implications about injury mechanisms. The American journal of sports medicine, 23(2), 233-239.

Fleisig, G. S., Escamilla, R. F., Andrews, J. R., Matsuo, T., Satterwhite, Y., & Barrentine, S. W. (1996). Kinematic and Kinetic Comparison between Baseball Pitching and Football Passing. *Journal of Applied Biomechanics*, *12*(2), 207–224. https://doi.org/10.1123/jab.12.2.207

Fleisig, G. S., Barrentine, S. W., Zheng, N., Escamilla, R. F., & Andrews, J. R. (1999). Kinematic and kinetic comparison of baseball pitching among various levels of development. *Journal of biomechanics*, *32*(12), 1371-1375.

Fleisig, G., Chu, Y., Weber, A., & Andrews, J. (2009). Variability in baseball pitching biomechanics among various levels of competition. *Sports Biomechanics*, *8*(1), 10–21. https://doi.org/10.1080/14763140802629958

Grantham WJ, Byram IR, Meadows MC, Ahmad CS. The Impact of Fatigue on the Kinematics of Collegiate Baseball Pitchers. *Orthopaedic Journal of Sports Medicine*. June 2014. doi:10.1177/2325967114537032

Gupta, V. (2019). Pose detection comparison: wrnchAI vs OpenPose. (April 15, 2020). https://www.learnopencv.com/pose-detectioncomparison-wrnchai-vs-openpose/

Ionescu, C., Papava, D., Olaru, V., & Sminchisescu, C. (2013). Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE transactions on pattern analysis and machine intelligence*, *36*(7), 1325-1339.

Kanko, R. M., Laende, E., Selbie, W. S., & Deluzio, K. J. (2021). Inter-session repeatability of markerless motion capture gait kinematics. *Journal of Biomechanics*, *121*, 110422.

McKinnon, C. D., Sonne, M. W., & Keir, P. J. (2020). Assessment of Joint Angle and Reach Envelope Demands Using a Video-Based Physical Demands Description Tool. *Human Factors*, 0018720820951349. https://doi.org/10.1177/0018720820951349

Mündermann, L., Corazza, S. & Andriacchi, T.P. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. J NeuroEngineering Rehabil 3, 6 (2006). https://doi.org/10.1186/1743-0003-3-6

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in Sports and Active Living*, *2*. https://doi.org/10.3389/fspor.2020.00050

Nicholls, R., Fleisig, G., Elliott, B., Lyman, S., & Osinski, E. (2003). Baseball: Accuracy of qualitative analysis for assessment of skilled baseball pitching technique. *Sports Biomechanics*, *2*(2), 213–226.

Qiao, S., Wang, Y., & Li, J. (2017). Real-time human gesture grading based on OpenPose. Proceedings of the 10th International Congress in Image and Signal Processing, Biomedical Engineering and Informatics (CISP-BMEI).

Richards, J. G. (1999). The measurement of human motion: A comparison of commercially available systems. *Human movement science*, *18*(5), 589-602.

Scarborough, D. M., Bassett, A. J., Mayer, L. W., & Berkson, E. M. (2020). Kinematic sequence patterns in the overhead baseball pitch. *Sports biomechanics*, *19*(5), 569-586.

Seroyer, Shane T., et al. "The kinetic chain in overhand pitching: its potential role for performance enhancement and injury prevention." *Sports health* 2.2 (2010): 135-146.

Stodden, D. F., Fleisig, G. S., McLean, S. P., & Andrews, J. R. (2005). Relationship of Biomechanical Factors to Baseball Pitching Velocity: Within Pitcher Variation. *Journal of Applied Biomechanics*, *21*(1), 44–56. https://doi.org/10.1123/jab.21.1.44

Tanaka, R., Takimoto, H., Yamasaki, T., & Higashi, A. (2018). Validity of time series kinematical data as measured by a markerless motion capture system on a flatland for gait assessment. *Journal of biomechanics*, *71*, 281-285.

Thewlis, D., Bishop, C., Daniell, N., & Paul, G. (2013). Next-generation low-cost motion capture systems can provide comparable spatial accuracy to high-end systems. *Journal of applied biomechanics*, *29*(1), 112-117.

Tian, Z. Z., Kyte, M. D., & Messer, C. J. (2002). Parallax error in video-image systems. *Journal of Transportation Engineering*, *128*(3), 218-223.

Vafadar, S., Skalli, W., Bonnet-Lebrun, A., Khalifé, M., Renaudin, M., Hamza, A., & Gajny, L. (2021). A novel dataset and deep learning-based approach for marker-less motion capture during gait. Gait & Posture, 86, 70-76.

Veeger, H.E.J. (2000). The position of the rotation center of the glenohumeral joint. *Journal of Biomechanics*, 33, 1711-1715.

Werner, S. L., Fleisig, G. S., Dillman, C. J., & Andrews, J. R. (1993). Biomechanics of the Elbow During Baseball Pitching. *Journal of Orthopaedic & Sports Physical Therapy*, *17*(6), 274–278. https://doi.org/10.2519/jospt.1993.17.6.274

Winter, D.A. (1990). Biomechanics and motor control of human movement. New York: Wiley Interscience.

Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M., D'Lima, D. D., Cristofolini, L., Witte, H., Schmid, O., Stokes, I., ., & Standardization and Terminology Committee of the International Society of Biomechanics. (2002). ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion--part I: ankle, hip, and spine. International Society of Biomechanics. *Journal of Biomechanics, 35*, 543–548.

Wu, G., van der Helm, F. C. T., Veeger, H. E. J. D., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X., Werner, F. W., Buchholz, B., ., & International Society of Biomechanics. (2005). 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. *Journal of Biomechanics, 38*, 981–992.

Yang J, Mann BJ, Guettler JH, et al. Risk-Prone Pitching Activities and Injuries in Youth Baseball: Findings From a National Sample. *The American Journal of Sports Medicine*. 2014;42(6):1456-1463. doi:10.1177/0363546514524699