Journal of Bodywork & Movement Therapies 40 (2024) 315–319

Contents lists available at ScienceDirect

Journal of Bodywork & Movement Therapies

journal homepage: www.elsevier.com/jbmt

Comparison of computational pose estimation models for joint angles with 3D motion capture

Rebecca I. Hamilton[a,*], Zornitza Glavcheva-Laleva[b], Md Imdadul Haque Milon[d], Yeshwin Anil[d], Jenny Williams[b], Peter Bishop[c], Catherine Holt[b]

[a] Centre for Trials Research, School of Medicine, Cardiff University, CF14 4YU, UK
[b] Musculoskeletal Biomechanics Research Facility, School of Engineering, Cardiff University, CF24 3AA, UK
[c] Agile Kinetic Ltd. Tramshed Tech, Griffin House, Griffin St, Newport, NP20 1GL, UK
[d] Cardiff Metropolitan University, Llandaff Campus, Western Ave, Cardiff, CF5 2YB, UK

ARTICLE INFO

Handling Editor: Dr Jerrilyn Cambron

1. Introduction

Human pose estimation is a growing technique and of particular benefit to clinicians and researchers working in musculoskeletal biomechanics interested in calculating human movement patterns (Stenum et al., 2021). The observation of a patient's joint angle during a movement can determine the effectiveness of a rehabilitation programme, risk of injury and other quality of life measures (Dos Santos et al., 2016) increasing the value of objectively assessing joint angle and joint range of motion (ROM) for allied health professionals. Where access to these sophisticated three-dimensional motion analysis (3DMA) systems are limited, the benefit of open-source tools that can be used in a variety of applications has a growing platform.

The established laboratory reference standard for joint angle/ROM calculations uses 3DMA technology, utilising 3D optical retroreflective marker-based systems and multiple video cameras (e.g., Qualisys, Vicon, Optitrack) (Keogh et al., 2019). The development of depth cameras has enabled marker-less solutions for joint angle measurement; now heavily used within 3DMA and rehabilitation research (Kanko et al., 2021). Though these tend to require specialist input, the development of trained two-dimensional (2D) pose estimation models makes objective human movement data collection more accessible, using small devices such as phone cameras (Halilaj et al., 2021). Human pose estimation localises body key points to accurately recognise the posture of individuals in an image or video (Munea et al., 2020).

Accuracy of various systems is likely to be assessed differently also limiting direct comparison. The accuracy of the Qualisys Track Manager (QTM) 3D laboratory reference standard uses a recommended residual error margin of < 1 mm from each camera perspective. Whereas pose estimation models often use mean per joint position error (MPJPE), utilising the mean of the distances between the estimated coordinate and the true coordinate, over each joint. This produces estimates for coordinate and skeletal position based on key points. With pose estimation resulting in an accepted reduced accuracy compared to laboratory-based 3DMA, MPJPE is reported to be accepted within a 20 mm error (Desmarais et al., 2021). The question remains as to whether the reduced accuracy of these systems can still give reliable ROM and joint angle estimates when compared to 3D laboratory standard systems given their different measurement methods and user applications.

The research objective of this exploratory study was to compare two human pose estimation models against the laboratory reference standard for the measurement of joint angle and ROM parameters, using basic comparative validation statistics with time matched video recordings. Comparisons were made between several different but simple movement activities that assess active ROM based on the recommendation from a clinical specialist in the field. The simplicity of the analysis demonstrates preliminary observations available for pose estimation examination and its suitability when restricted by resource and time. However, there remains further potential for in depth analysis to determine specific model differences that arise with particular movements types and situations to determine more granular levels of accuracy between models.

* Corresponding author.
E-mail address: HamiltonR@cardiff.ac.uk (R.I. Hamilton).

https://doi.org/10.1016/j.jbmt.2024.04.033
Received 16 May 2023; Received in revised form 15 March 2024; Accepted 9 April 2024
Available online 16 April 2024
1360-8592/

2. Methods

2.1. Participant data collection

Twenty two healthy volunteers with no history of injury or other lower and upper limb pathology, participated in this exploratory study to compare pose estimation model outputs with a QTM 3DMA marker-based laboratory system.

After the study was ethically approved by a Research Ethics Committee, all subjects were provided with an information sheet and gave written informed consent on the day of data collection. Participants were recruited via social media and word of mouth advertisement within and around the Musculoskeletal Biomechanics Research Facility (MSKBRF) at Cardiff University, School of Engineering. Participants were asked to wear loose or compressed shorts, a loose vest/T-shirt, and comfortable shoes/trainers before performing the following set of activities: sit-to-stand for five repetitions and one repetition of seated right knee extension, standing right elbow flexion and prone left knee flexion. Activities were recorded using marker-based 3DMA (Qualisys, Sweden) with 12 Oqus 700 + infrared cameras and 2 Oqus 210c video cameras for 2D image collection, capturing at 100Hz and 24Hz respectively. The infrared and video cameras were synchronised using the QTM trigger module which sends out a TTL pulse to start and stop recording. Data collection took place in a Motion Analysis Laboratory in the MSKBRF. A modified upper and lower body marker placement protocol, with 12 retro-reflective markers, was used for the data collection to represent the target points for the pose estimation software (Fig. 1). The markers were placed at specific anatomical locations on the right and left acromion, lateral elbow epicondyles, wrist pisiform bone, greater trochanter, lateral knee epicondyles, and lateral malleoli.

The subjects were video recorded in the sagittal plane as they performed the activities to translate the outputs into the pose estimation models. The video camera was positioned 4.2m from the subject for sit-to-stand, seated right knee extension and standing right elbow flexion activities, and 4.8m from the subject for prone left knee flexion activity.

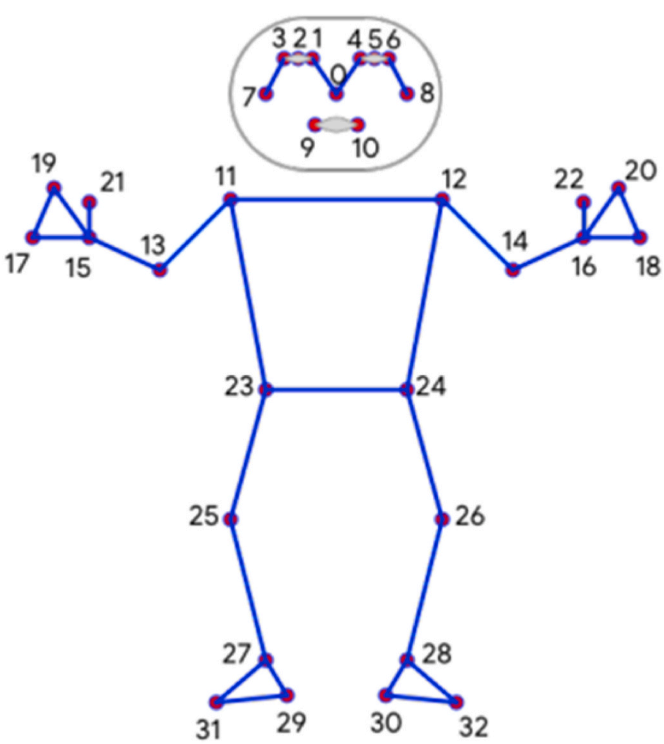

Fig. 1. 33 key point topology of the MediaPipe model (Bazarevsky et al., 2020).

2.2. Kinematic outputs

The coordinates of the markers placed at the acromion, lateral elbow epicondyles, and wrist pisiform and at the greater trochanter, lateral femoral epicondyles, and lateral malleoli were used to calculate elbow and knee joint angles, as kinematic variables, respectively, within QTM.

Two algorithms, HRNet and MediaPipe were used for comparisons as current open-source benchmark 2D pose estimation models (Sun et al., 2019; Bazarevsky et al., 2020). MediaPipe uses a combination of heatmap, offset and regression approaches to train the network. The heatmaps and offsets for training are then removed when reading in data, making it a light-weight architecture. An encoder-decoder network followed by regression encoder network then generates the 33 key points on the body (Fig. 1). The HRNet (Sun et al., 2019) uses multi-resolution subnetworks enabling the maintenance of high-resolution features across the whole process, and contributes to more accurate and spatially precise heatmaps. Key points are predicted from the high-resolution features using a key point detection dataset to predict 17 key points trained on Common Objects in Context (Lin et al., 2014). Using the predicted key points from the MediaPipe and HRNet models, metrics for musculoskeletal kinematic analysis were implemented. Extracted key points provided by the pose estimation networks were used to identify joints and calculate joint angles.

Joint ROM was defined as the difference between maximum and minimum joint angles, calculated using coordinate positions from the detected key points, in radians then converted into degrees.

2.3. Output comparisons and statistical analysis

Results across the three systems were matched using the resultant time frames and time stamps. Since the video and infrared cameras were synchronised, time stamps could be calculated with respect to the starting frame for both QTM and the pose estimation data derived from the video camera. The joint angles calculated were defined by key points with the marker placement and recorded throughout the exercise as a time series for each participant and matched with respect to the starting frame. Thus, the three systems had joint angle results at matched time stamps and could then be compared across each other for Coefficient of Variation (CoV) and Standard Deviation (SD) statistics. The mean of the CoV and SD values for each participant and each exercise were presented in summary data (Table 1).

Statistical comparisons for joint angle and ROM as a discrete

Table 1
Coefficient of Variation (CoV) and Standard Deviation (SD) for joint angle time series data across the different systems (MoCap, MediaPipe, HRNet) for each activity, including the key points utilised to determine the joint angles.

Activity	3D key points used	Mean MoCap, MediaPipe & HRNet CoV (±SD)	Mean MoCap & MediaPipe CoV (±SD)	Mean MoCap & HRNet CoV (±SD)
Seated Knee Extension Flexion (R)	Right ankle, knee, and hip	4.638 ± 1.95	**4.6** ± 2.15	4.69 ± 2.72
Prone Knee Flexion Extension (L)	Left ankle, knee, and hip	5.535 ± 3.82	**4.86** ± 2.3	5.12 ± 3.86
Elbow Flexion Extension (R)	Right wrist, elbow, and shoulder	5.335 ± 2.24	5.81 ± 2.22	**5.45** ± 3.15
Sit to Stand Knee (L)	Left ankle, knee, and hip	4.322 ± 1.53	4.57 ± 2.28	**4.29** ± 1.7
Sit to Stand Knee (R)	Right ankle, knee, and hip	4.005 ± 1.39	4.65 ± 1.79	**3.66** ± 1.62

parameter were computed using CoV and Intra-Class Correlation Coefficient (ICC). Data descriptions and statistical comparisons were computed with Microsoft Excel and SPSS. CoV comparisons were calculated for ROM data with SD values and ICC to measure the level of agreement consistency as a correlation, presented with statistical significance (Table 2). ICC results are categorised as follows: above 0.9 = excellent reliability, 0.75–0.9 good, 0.5-.75 moderate and below 0.5 = poor reliability (Koo et al., 2016) and strengthened with a significant p value < 0.05 using an alpha confidence interval of 95%.

3. Results

The data represents the results from twenty-two healthy population volunteers (Female n = 16, Male n = 6, Age mean = 36.9 ± 11.6, Weight (kg) mean = 75.9 ± 15.8, Height (m) mean = 1.7 ± 0.1, BMI (kg/m^2) mean = 26.1 ± 3.8).

3.1. Descriptive statistics

Table 1 displays descriptive results for joint angle calculated by the three different systems: 3DMA joint angle, vision-based MediaPipe and vision-based HRNet. The vision-based model with the best comparison score when assessed against marker based joint angle system is denoted in **bold.** Table 1 demonstrates MediaPipe as having the lower variation value for seated and prone knee extension-flexion exercises when compared to 3DMA, whereas HRNet displays lower variation for the other 3 activities. All resulted in low summary variations but with high deviations. A participant example of joint angle data for the 5 outputs of 4 movement activities is demonstrated in Fig. 2 (note these do not represent summary data but rather example data from one participant).

3.2. Comparison statistics

CoV values for the ROM comparisons in Table 2 demonstrate similar variation results to the joint angle time series results in Table 1 with MediaPipe demonstrating lower variation values for seated and prone knee extension-flexion, only including left knee sit-to-stand exercise also. The HRNet displays lower variation for the other two exercises. Each exercise displayed strong and significant correlations for both models when comparing to 3DMA joint angle apart from left knee sit-to-stand which also has a higher CoV for both model comparisons.

4. Discussion

When investigating the use of pose estimation model outputs, their comparison to 3DMA outputs can be explored in a variety of manners.

Table 2
Range of motion statistical comparison summary table utilising Intra-Class Correlation Coefficient (ICC) with significant correlation denoted by *.

Activity	MoCap vs MediaPipe CoV (% ±SD)	MoCap vs HRNet CoV (% ±SD)	MoCap vs MediaPipe ICC (with sig.)	MoCap vs HRNet ICC (with sig.)
Seated Knee Extension (R)	**8.02 ± 3.8**	9.73 ± 5.21	**0.95 (p < 0.001)***	0.87 (p < 0.001)*
Prone Knee Flexion Extension (L)	**7.98 ± 7.08**	10.02 ± 8.66	**0.81 (p < 0.001)***	0.63 (p = 0.021)*
Elbow Flexion Extension (R)	4.05 ± 2.97	**3.25 ± 2.96**	0.92 (p < 0.001)*	**0.94 (p < 0.001)***
Sit to Stand Knee (L)	**11.68 ± 9.66**	11.71 ± 7.23	0.41 (p = 0.138)	0.41 (p = 0.134)
Sit to Stand Knee (R)	11.66 ± 3.7	**10.53 ± 4.96**	**0.83 (p < 0.001)***	0.82 (p < 0.001)*

Minimal but valuable comparisons can demonstrate quick results to help decipher the validity of an open-source model with simple movements in a reduced setting. This could be similar to a clinic or treatment room setting.

The time series movement summary results for both models, display similar values when utilising CoV as a variation statistic. With all CoV results under 10% (Table 1), this implies that both models compute reliable joint angle time series data based on previous literature result standards utilising low variation as a measure (Dos Santos et al., 2016). Equally similar low variation results are seen in more established open source models such as OpenCap, when comparing both joint and angles and kinetic force data with 3DMA methods (Ulrich et al., 2023). Despite low CoV values for both models, dispersion values remain high for all findings, thus reducing their reliability, and this should be considered within the consistency agreement interpretation.

When comparing ROM as a discrete parameter, CoV values remain low, though not all below 10% such as joint angle results (Table 2). The two activities for sit-to-stand result in higher variation and dispersion values when compared to the other activities, implying that the pose estimation outputs for this activity are not sufficiently developed to be reliably comparable to 3DMA. This may be due to larger degrees of motion resulting in a higher likelihood of errors within the key point predictions, or specific movements obstructing the camera image. This could then affect the maximum and minimum angle distinctions and resultant ROM measure. Other similar models that struggle with complexities (Resnet and CPN) in movements are reported, such as heatmap loss (Sun et al., 2018), reconstruction loss (Xu et al., 2019) and motion loss (Wang et al., 2020) and may benefit from training on larger datasets with further complex movements or potentially use of threshold metrics rather than utilising accuracy comparisons.

The weaker comparison results for sit-to-stand activities are reinforced by outcomes displayed for ICC values. This determines the correlative similarity between the data, with good-to-excellent correlations for the three other activities (supported by level of significance determined), and lower correlation values for sit-to-stand results. The difference in correlation values between the two pose estimation models are so minor that it should be concluded that either of them would determine strong to excellent validity compared to 3DMA. Since sit-to-stand is one of the more prominent activities within clinical gait analysis and rehabilitation monitoring for populations such as elderly (Smith et al., 2020) and cerebral palsy (Apoorva et al., 2018), it is crucial to ensure this activity is associated with good-excellent degree of accuracy and reliability. Other models have shown greater reliability for this metric (Kidziński et al., 2020) and compared across clinically significant measures as a means of clinical setting validation. This is a crucial element to test the reliability of the tool to produce a clinically measurable difference and is required for the tools analysed in this paper if intended to be used in a clinical setting.

Though many of the results display similarities across the two pose estimation models when compared to 3DMA, certain activities and parameter outputs are marginally better for individual models. Exploration of the data reveals that seated knee extension/flexion and prone knee flexion/extension produce better results with MediaPipe, and elbow flexion/extension with HRNet, however, the differences between these models is minor using the applied variation and correlation measures. Both models show strong levels of validity when compared to the 3DMA laboratory standard, when based on a small cohort and simple descriptive statistics as a means of validity comparison. Development of further parameter outputs such as movement ROM, angular acceleration, angular velocity, movement symmetry, and movement consistency, with larger data sets will strengthen use reliability within a clinical field and allow for remote assessments.

Where these activities would generally be analysed in a more sophisticated setting, such models could provide transformational tools with a large degree of impact when used in more remote settings, given the practical difficulty of rehabilitation testing in laboratories. The need

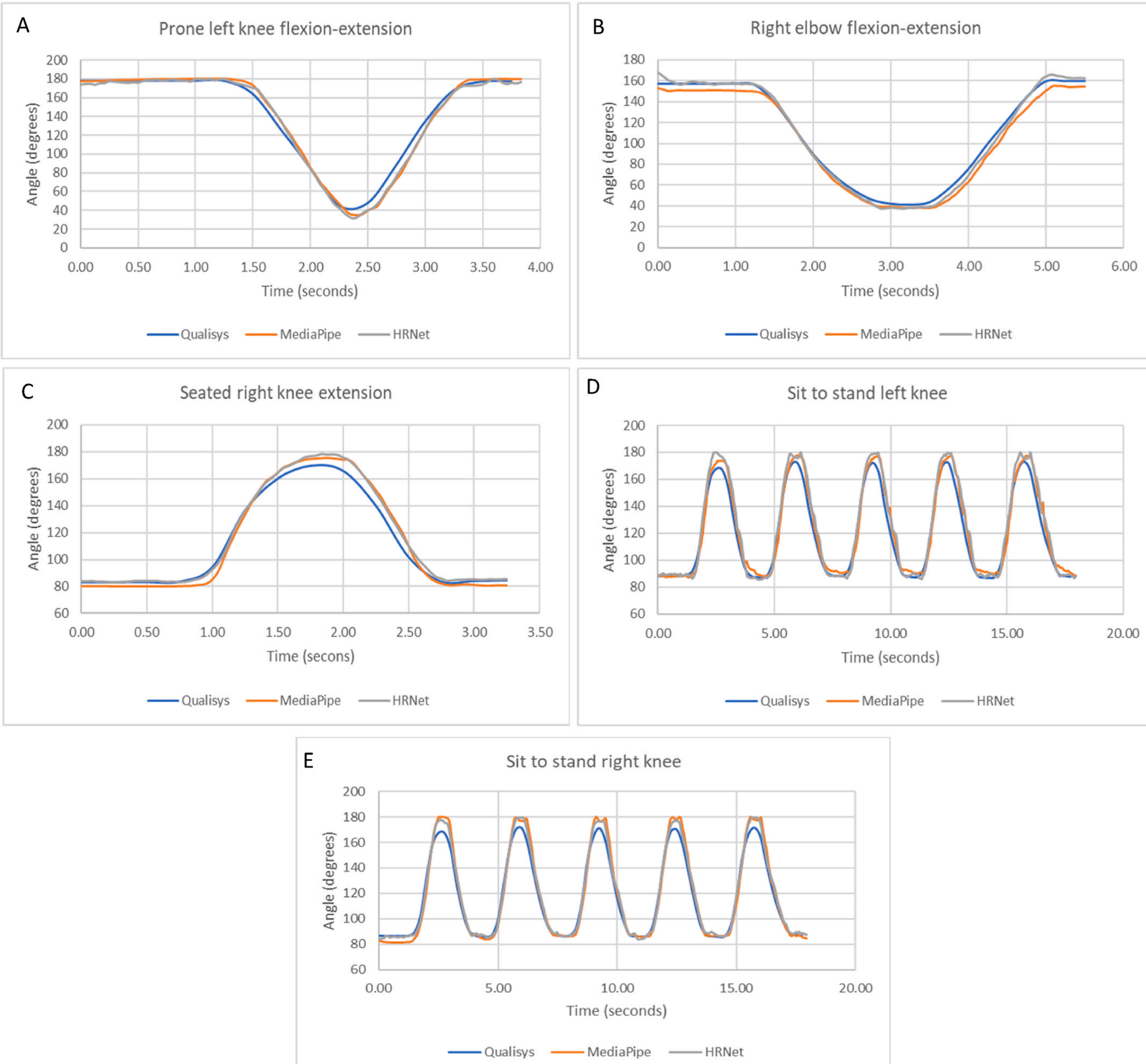

Fig. 2. Demonstration of joint angle similarity between the 3 models compared for one example participant results for the 5 outputs of 4 activities of daily living recorded (A – Prone left knee flexion extension, B – right elbow flexion extension, C – seated right knee extension, D – Sit to stand left knee, E – sit to stand right knee).

for "real world testing" and validation methods within these settings is increasingly apparent in the literature (Weygers et al., 2020; Stenum et al., 2021) due to known adaptations that people make when in a laboratory testing setting. Increasing pilot and validation data for such remote tools, provides opportunities for clinicians and patients to rely on remote tools to the same extent as laboratory or clinically based methods.

The need and desire to utilise pose estimation tools within clinical and rehabilitative settings is driven by the potential for fast and practical access to objective measures. Though reduced accuracy in comparison to 3DMA, this may still provide greater reliability when compared to subjective or self-reported measures. It also promises access to much larger scale human-based movement research using less resources with lower levels of required expertise, allowing growth of movement-based data sharing and dataset development (Ulrich et al., 2023). This will be critical for development and possibly stratification of patients in musculoskeletal conditions, such as osteoarthritis, where this data is simply not well accessed (Evans et al., 2022). Therefore, a degree of trade-off should be accepted in terms of reliability depending on the model use and its purpose. Particularly, threshold metrics may help to decipher what trade-off is acceptable depending on the setting.

Further measure comparisons that are used within pose estimation algorithm tools such as MPJPE, angular velocity, angular acceleration, movement symmetry, movement consistency as well as threshold metrics (reported to accurately identify errors specifically for joint detection) could provide greater granularity of results for movement detection, particularly for analysis of a range of complex movements (Desmarais et al., 2021).

This exploratory study has several limitations. The sample size and range of population (twenty-two participants, BMI 26.1 ± 3.8 and ≥60 years of age) are restrictive and should be considered in data interpretation. The data was collected on a general and healthy population and therefore does not represent many population groups that would be utilising these tools for rehabilitative exercises. These exploratory findings should be considered for future studies developed to achieve more relevant results for population groups with musculoskeletal disorders. The use of a more comprehensive marker set will allow for the limitations created when collecting motion data from a patient population and will enhance results further.

5. Conclusion

The results demonstrate that these 2D pose estimation models that are easy to use tools for remote rehabilitative purposes, can be compared and validated with 3DMA in a simple analysis to provide valuable insight. The results are stronger in flexion/extension movements compared to full sit-to-stand movements, however, this is based on a small selection of movements. The low variation statistics and strong correlations demonstrate that they provide a comparative standard to a reduced 3DMA analysis, depending on the context to which the models will be used. For many rehabilitative applications, these quick and easy to access tools meet the need to provide valuable, objective, movement-based information where subjective or self-reported measures may be the only alternative.

CRediT authorship contribution statement

Rebecca I. Hamilton: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Zornitza Glavcheva-Laleva:** Project administration, Methodology, Investigation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Md Imdadul Haque Milon:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis, Data curation, Validation, Visualization, Writing – original draft. **Yeshwin Anil:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis, Data curation, Validation, Visualization, Writing – original draft. **Jenny Williams:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Peter Bishop:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Catherine Holt:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors confirm that the content of the manuscript has not been published or submitted for publication in another journal but has been presented as exploratory data and protocol testing at various meetings.

The work has been a collaboration with Cardiff University and industry partners (Agile Kinetic Ltd.) who have commercial interest in pose estimation products.

All data collection and analysis were carried out by university research and technical staff with no commercial interest in these products.

This research was funded by Accelerate Wales, Life Sciences Hub Programme from Welsh Government European Regional Devlopment Fund (PR-0345).

There are no known conflicts of interest associated with this paper.

References

Apoorva, R., Kidziński, Ł., McGlaughlin, A.s., Hicks, J., Delp, S.L., Schwartz, M., 2018. Estimating the effect size of surgery to improve walking in children with cerebral palsy from retrospective observational clinical data. Sci. Rep. 8 (16344) https://doi.org/10.1038/s41598-018-33962-2.

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., Grundmann, M., 2020. BlazePose: On-Device Real-Time Body Pose Tracking. https://doi.org/10.48550/arxiv.2006.10204.

Desmarais, Y., Mottet, D., Slangen, P., Montesinos, P., 2021. A review of 3D human pose estimation algorithms for markerless motion capture. Comput. Vis. Image Understand. 212, 103275 https://doi.org/10.1016/j.cviu.2021.103275.

Dos Santos, R.A., Derhon, V., Brandalize, M., Brandalize, D., Rossi, L.P., 2016. Evaluation of knee range of motion: correlation between measurements using a universal goniometer and a smartphone goniometric application. J. Bodyw. Mov. Ther. 21 (3), 699–703. https://doi.org/10.1016/j.jbmt.2016.11.008.

Evans, J., Hamilton, R.I., Biggs, P., Holt, C., Elliott, M., 2022. Data sharing across osteoarthritis research groups and disciplines: opportunities and challenges. Osteoarthritis and Cartilage Open 4 (1). https://doi.org/10.1016/j.ocarto.2022.100236.

Halilaj, E., Shin, S., Rapp, E., Xiang, D., 2021. American society of biomechanics early career achievement award 2020: toward portable and modular biomechanics labs: how video and IMU fusion will change gait analysis. J. Biomech. 129, 110650 https://doi.org/10.1016/j.jbiomech.2021.110650.

Kanko, R.M., Laende, E.K., Davis, E.M., Selbie, W.S., Deluzio, K.J., 2021. Concurrent assessment of gait kinematics using marker-based and markerless motion capture. J. Biomech. 127, 110665 https://doi.org/10.1016/j.jbiomech.2021.110665.

Keogh, J.W.L., Cox, A., Anderson, S., Liew, B., Olsen, A., Schram, B., Furness, J., 2019. Reliability and validity of clinically accessible smartphone applications to measure joint range of motion: a systematic review. PLoS One 14 (5), e0215806. https://doi.org/10.1371/journal.pone.0215806.

Kidziński, Ł., Yang, B., Hicks, J., Rajagopal, A., Delp, S.L., Schwartz, M.H., 2020. Deep neural networks enable quantitative movement analysis using single-camera videos. Nat. Commun. 11 (4054) https://doi.org/10.1038/s41467-020-17807.

Koo, T., Li, Mae, 2016. A gudeline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractice Medicine 15 (2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Lin, T.-Y., et al., 2014. Microsoft COCO: Common Objects in Context. https://doi.org/10.48550/arxiv.1405.0312.

Munea, T.L., Jembre, Y.Z., Weldegebriel, H.T., Chen, L., Huang, C., Yang, C., 2020. The progress of human pose estimation: a survey and taxonomy of models applied in 2D human pose estimation. IEEE Access 8, 133330–133348. https://doi.org/10.1109/ACCESS.2020.3010248.

Smith, S.H.L., Reilly, P., Bull, A., 2020. A musculoskeletal modelling approach to explain sit-to-stand difficulties in older people due to changes in muscle recruitment and movement strategies. J. Biomech. 98 (109451) https://doi.org/10.1016/j.jbiomech.2019.109451.

Stenum, J., Cherry-Allen, K.M., Pyles, C.O., Reetzke, R.D., Vignos, M.F., Roemmich, R.T., 2021. Applications of pose estimation in human health and performance across the lifespan. Sensors 21 (21), 7315. https://doi.org/10.3390/s21217315.

Sun, X., Xiao, B.Wei, Flk Liang, S., Wei, Y., 2018. Integral Human Pose Regression. https://doi.org/10.48550/arXiv.1711.08229.

Sun, K., Xiao, B., Liu, D., Wang, J., 2019. Deep High-Resolution Representation Learning for Human Pose Estimation. https://doi.org/10.48550/arxiv.1902.09212.

Ulrich, S., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A., Hicks, J., Delp, S., 2023. OpenCap: human movement dynamics from smartphone videos. PLoS Comput. Biol. 19 (10) https://doi.org/10.1371/journal.pcbi.1011462.

Wang, J., Yan, S., Xiong, Y., Lin, D., 2020. Motion guided 3D pose estimation from videos. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, j.M. (Eds.), Computer Vision - ECCV 2020, vol. 12358. Springer International Publishing, Cnam, pp. 764–780. https://doi.org/10.1007/978-3-030-58601-0_45.

Weygers, I., Kok, M., Konings, M.Hallez, De Vroey, H., Claeys, K., 2020. Inertial sensor-based lower limb joint kinematics: a methodological systematic review. Sensors 20 (673). https://doi.org/10.3390/s20030673.

Xu, Y., Zhu, S.C., Tung, T., 2019. DenseRaC: joint 3D pose and shape estimstion by dense render-and-compare. arXiv: 1910, 00116.