Computers in Biology and Medicine 171 (2024) 108101

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

On the reliability of single-camera markerless systems for overground gait monitoring

Michele Boldo [a], Roberto Di Marco [b,*], Enrico Martini [a], Mauro Nardon [c], Matteo Bertucco [c], Nicola Bombieri [b]

[a] Department of Computer Science, University of Verona, Strada Le Grazie, 15, Verona, 37134, Italy
[b] Department of Engineering for Innovation Medicine, University of Verona, Strada Le Grazie, 15, Verona, 37134, Italy
[c] Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Via Casorati, 43, Verona, 37131, Italy

ARTICLE INFO

Keywords:
Motion analysis
Human pose estimation
Human joint kinematics
Clinical evaluation
Reference data
Telemedicine

ABSTRACT

Background and Objective: Motion analysis is crucial for effective and timely rehabilitative interventions on people with motor disorders. Conventional marker-based (MB) gait analysis is highly time-consuming and calls for expensive equipment, dedicated facilities and personnel. Markerless (ML) systems may pave the way to less demanding gait monitoring, also in unsupervised environments (i.e., in telemedicine). However,scepticism on clinical usability of relevant outcome measures has hampered its use. ML is normally used to analyse treadmill walking, which is significantly different from the more physiological overground walking. This study aims to provide end-users with instructions on using a single-camera markerless system to obtain reliable motion data from overground walking, while clinicians will be instructed on the reliability of obtained quantities.
Methods: The study compares kinematics obtained from ML systems to those concurrently obtained from marker-based systems, considering different stride counts and subject positioning within the capture volume.
Results: The findings suggest that five straight walking trials are sufficient for collecting reliable kinematics with ML systems. Precision on joint kinematics decreased at the boundary of the capture volume. Excellent correlation was found between ML and MB systems for hip and knee angles ($0.92 < R^2 < 0.96$), with slightly lower correlations observed for ankle plantar-dorsiflexion. The Bland–Altman analysis indicated the largest bias for hip flexion/extension ([0.2°, 10.9°]) and the smallest for knee joint ([0.1°, 0.8°]) when comparing MB-PiG and MB-JC approaches. For MB-JC vs. ML-JC comparison, the largest bias was for the ankle joint ([1.2°, 11.8°]), while the smallest was for the hip joint ([0.2°, 7.3°]).
Conclusion: Single-camera markerless motion capture systems have great potential in assessing human joint kinematics during overground walking. Clinicians can confidently rely on estimated joint kinematics while walking, enabling personalized interventions and improving accessibility to remote evaluation and rehabilitation services, as long as: (i) the camera is positioned to capture someone walking back and forth at least five times with good visibility of the entire body silhouette; (ii) the walking path is at least 2 m long; and (iii) images captured at the boundaries of the camera image plane should be discarded.

1. Introduction

Ageing and various pathologies cause depletion of individuals' balance and motor abilities, representing a major health issue that leads to high healthcare costs [1,2]. Effective and timely interventions on individuals with balance and motor disorders are crucial to restore acceptable motor abilities, improve people's quality of life and relieve healthcare facilities from cost-consuming rehabilitative interventions, which also impact patients' economy [3].

Motion analysis plays an important role to this purpose, but it is normally performed via stereophotogrammetry, which uses bulky and expensive systems, calling for: (i) extensive subject preparation; (ii) a team of expert operators; (iii) the subject to walk barefoot and nearly undressed; (iv) a dedicated environment [4–7]; and (v) extensive pre-processing procedures (i.e., trajectories' labelling, gap filling, smoothing, etc.). These aspects hamper the use of such systems to assess individuals' motion abilities, with the consequent lack of information on their progress during the rehabilitative intervention.

* Corresponding author.
E-mail addresses: michele.boldo@univr.it (M. Boldo), roberto.dimarco@univr.it (R. Di Marco), enrico.martini@univr.it (E. Martini), mauro.nardon@univr.it (M. Nardon), matteo.bertucco@univr.it (M. Bertucco), nicola.bombieri@univr.it (N. Bombieri).

https://doi.org/10.1016/j.compbiomed.2024.108101
Received 9 June 2023; Received in revised form 16 January 2024; Accepted 4 February 2024
Available online 6 February 2024

Valid alternatives to monitor individuals' motion skills, which the scientific community is increasing interest in, are the deep learning techniques applied to motion and pattern recognition [8], and the markerless (ML) motion capture [9,10]. This latter approach would overcome most of the above-listed limitations, as they could be: (i) used both in clinical settings and unsupervised environments and with individuals potentially performing daily life activities [10]; (ii) used with no need for specialized operators [10]; (iii) low-cost [11]; (iv) free from errors associated with placing markers on individuals' skin (such as soft-tissue artefact and landmarks misrecognition) [12–14]; (v) used both for clinical evaluations and to check adherence to therapies [15,16] and (vi) potentially automatic, with no need of manual pre-processing typical of marker-based systems. These potential strengths not only enable the provision of clinicians with reliable data about human gait during telemedicine applications, but also facilitate motion screening in more familiar environments, thereby alleviating both patients and the healthcare system from costs associated with the utilization of clinical facilities [17,18].

Many ML systems have been proposed and validated. They are normally based on convolutional neural networks (CNN) to track feature points (the *keypoints*) on images captured through a standard digital camera [19]. Several authors studied the appropriateness of using markerless systems for clinical gait analysis [20–22]. However, many proposed systems implemented multi-camera solutions to study treadmill walking. Besides being expensive and, thus, restricting its use again to clinical facilities, treadmills should be cautiously considered when performing motion analysis of gait. Indeed, treadmill and overground walking are entirely different tasks from a neurophysiological point of view. The former calls for stabilizing the centre of mass (COM) while moving the lower limbs to answer the treadmill action, i.e., it aims to find COM stability. The latter consists of moving the COM forward by cyclically shifting the body weight from a supporting limb to the other: i.e., it aims to manage a sequence of balanced and unbalanced COM positions.

Researchers are working to develop more accurate markerless systems to be implemented in many different contexts, including telemedicine [17,18]. As said, motion analysis is crucial to guide interventions and study their effectiveness on patients, and telemedicine represents a valuable alternative to in-center evaluation and rehabilitation, as it: (i) reduces the cost of intervention for both the healthcare system and users [18,23,24]; (ii) grants access to therapies to people living far from health centers and with reduced mobility [25]; (iii) guarantees care continuity, either after discharge, or to prevent fragile individuals from contracting healthcare-associated infections (as during the COVID-19 pandemic [26,27]), with no need for the patients to travel back to health centers; (iv) boosts adherence to therapies [15,16]; and (v) increases individuals' engagement to the intervention [28]. An example could be a motor and cognitive therapies delivered via interactive Information and Communication Technologies, with patients and health professionals far apart [18].

Despite its scientific and practical soundness, telemedicine is not yet widely disseminated [25]. One possible reason could be the lack of clear instructions on how to use measurement systems outside from laboratories while still obtaining reliable results. Bringing this closer to motion analysis, guidelines should include answers to the following questions:

- **RQ1** How many times should a user walk back and forth in front of the ML system in order to collect sufficient data to obtain reliable kinematics? [29]
- **RQ2** Is there a part of the capture volume to be preferred to record motion data within? [5,30]
- **RQ3** **(a)** Is ML kinematics reliable enough for clinical use? [10,29] – **(b)** Is it possible to consider data gathered from the side of the body partially occluded to the camera objective? [17]

It is worth underlying that there is a trade-off between usability and accuracy of ML systems [31], particularly when considering single-camera rather than multi-camera systems, as these are more suitable to expert audience as they call for complex calibration procedures to be performed [10].

The aim of this work is to instruct: (i) end-users of a single-camera markerless system [17] on how to collect reliable data to be shared with clinicians, via tele-health networks, for monitoring of their walking abilities; and (ii) clinicians on which outcome variables are reliable. Ultimately, the methodology proposed within the present study could be considered as a generic assessment framework to test different markerless systems either composed by different single-cameras, or multi-camera systems and systems employing different pose estimators. Indeed, given an ML system, the employment of the proposed methodology provides information on how to use the system itself to obtain reliable measures, and to what extent the assessment of joint kinematics with the ML system is different from the kinematics obtained with the MB system. Materials and methods, including the data flow from capture to analysis and statistics, are described in Section 2. Results given in Section 3 and in the Supplemental Materials, and the limitation of the study are discussed in Section 4.

2. Materials and methods

To pursue the above mentioned objectives, we implemented the methodology described in Fig. 1 and detailed in the paragraphs 2.2, 2.3, 2.4, 2.5. Participants were asked to walk self-paced overground while motion data were synchronously collected both with a marker-based (MB) and a markerless (ML) systems, with the latter consisting of an off-the-shelf RGB camera and a pose estimator. A Python script interacts with the RGB sensor, collecting the images at a predetermined frequency. The RGB images are processed by another Python script that implements a neural network-based human pose estimation framework, which extrapolates the joint positions in 2D.

Kinematics calculated from MB and ML data were compared, varying the number of considered strides and capture volume to test the hypotheses that (i) the higher the stride number and (ii) the more the capture volume was centred with respect to the camera objective, the more precise the kinematics.

2.1. Participants and ethics statement

Eight healthy adults participated in the study (6 males and 2 females, 28.0 ± 3.7 years old, body mass 68.1 ± 14.9 kg, height 1.73 ± 0.12 m). None of the participants had a previous history of traumas or neuropathies. All participants provided written informed consent before taking part in the experimental procedures. The study protocol was performed in accordance with the Declaration of Helsinki and was approved by the Ethics Committee for Human Research of the University of Verona (Approval number 08.R3/2021).

2.2. Equipment and data collection

2.2.1. Marker-based motion capture

Prior to data collection, an expert human movement scientist collected the anthropometry from the participants, as prescribed by the guidelines of the full-body Plug-in-Gait (PiG) protocol for marker-placement (which is the Vicon's commercial version of the Davis protocol [32]): height, body-mass, distance between the anterior superior iliac spines, leg length, medio-lateral inter-epicondyles distance of the knee, medio-lateral inter-malleoli distance, shoulder width (i.e., the distance between the acromion marker and the shoulder joint center), inter-epicondyles distance of the elbow and hand thickness. The same technician equipped the participants with 39 retro-reflective skin-mounted markers placed according to the PiG protocol. The floor was marked with 11 markers to easily map the relative position of the

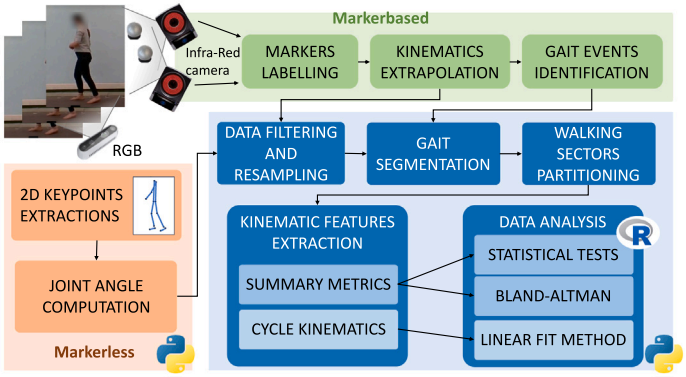

Fig. 1. Overview of the proposed data collection architecture. Participants' motion is concurrently recorded by infra-red cameras for marker-based gait analysis (green), and a RGB sensor for markerless approach (orange). Data gathered with the two approaches were stored, pre-processed, filtered and analysed (blue).

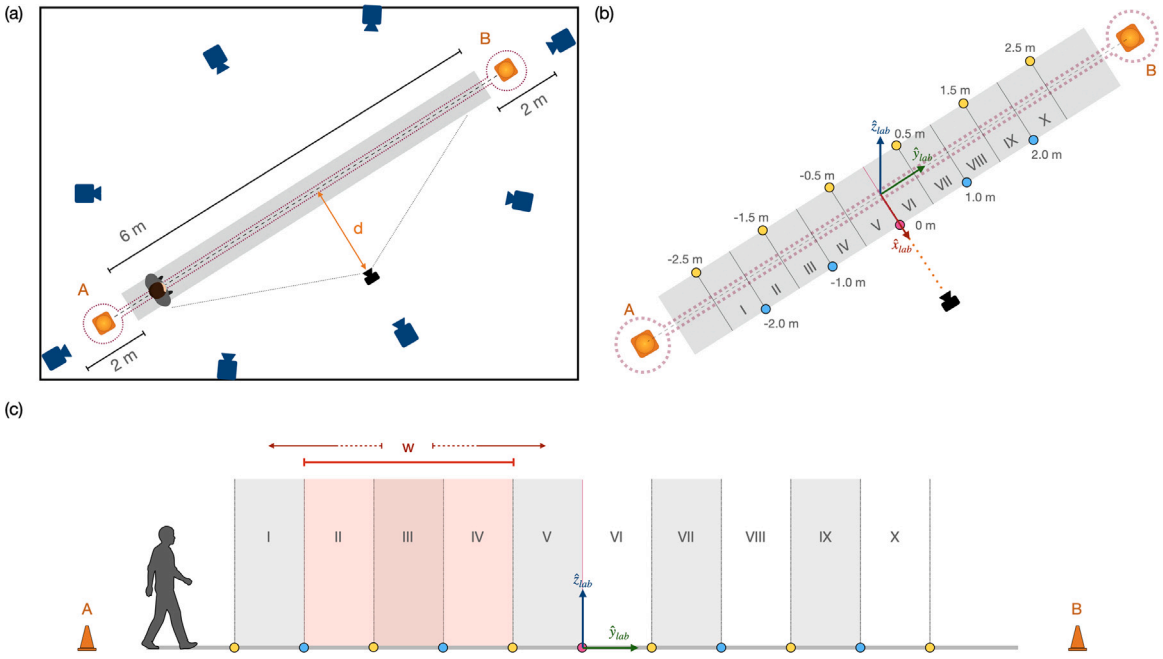

Fig. 2. Camera and walking path setup. **(a) Camera setup:** blue cameras are those pertaining to the marker-based (MB) system; the markerless RGB-D camera is in black and placed at a distance d from the walking path mid-line. **(b) Walking path — top view.** Participants walked back and forth on a 6 m pathway (highlighted in gray) within the capture volume of the MB system, following the path highlighted by the purple dotted line from cone A to cone B (in orange). The $\hat{x}, \hat{y}, \hat{z}_{lab}$ represent the global laboratory reference frame for the MB system. The RGB-D camera is aligned with the x-axis of this coordinate system (camera global coordinates on the floor are $x_{lab} = d$ and $y_{lab} = 0$). The purple circle pictures the marker placed along the $\hat{x}_{lab}$ direction. Eleven more markers (yellow and light blue circles) are placed on the walking path to mark 0.5 m width sectors (numbered with Roman numeral from I to X). **(c) Walking path — sagittal view.** The sliding light-red window (width $w = 1.5$ m) highlights the progressive selection of sectors of vision.

participants while walking with respect to the drawn walking path and the ML camera. This allowed defining 11 sectors of vision, identified with Roman numeral from I to X, each of 0.5 m width (Fig. 2b-c).

Three-dimensional marker time-histories were collected through a MB motion capture system (sampling rate: 120 Hz), which consisted of eight infrared emitting cameras (MX 13, Vicon, Oxford, UK - Fig. 2a) and using the proprietary software Vicon Nexus 2.14.

2.2.2. Markerless motion capture

Motion data were concurrently collected with a single-camera ML system, consisting of an Intel RealSense D415 and a Human Pose Estimation (HPE) framework. The Intel RealSense D415 was equipped with a RGB-D camera, which is able to provide both conventional RGB images and depth information calculated through active stereoscopic processing. To the purpose of the present study, data from the depth sensor was not used to feed the HPE framework with.

The RealSense was placed perpendicular to the walking path (Fig. 2a-b): $x_{lab} = d$, with the distance d being equal to 5 m for all the participants. This choice was made to consider a capture volume that fits the 6 m walking track (Fig. 2a).

The resolution was set to 848 × 480 pixels for both RGB and Depth cameras, with sampling rate equal to 60 fps and Field of View (FOV) 69° × 42°.

The HPE framework was used to extrapolate the human pose into the 3D space. More specifically, a HPE framework is able to detect and track specific keypoints, such as: shoulder, hip, knee and ankle joint centers. The chosen framework for this research was OpenPose [33], which is considered as the state of the art in terms of accuracy, but it is computationally heavier than others. Precisely, we considered the convolutional neural network (CNN) Body_25 model (25 keypoints to represent the human body), trained on COCO [34] (with its latest version: [35]) and MPII datasets [36].

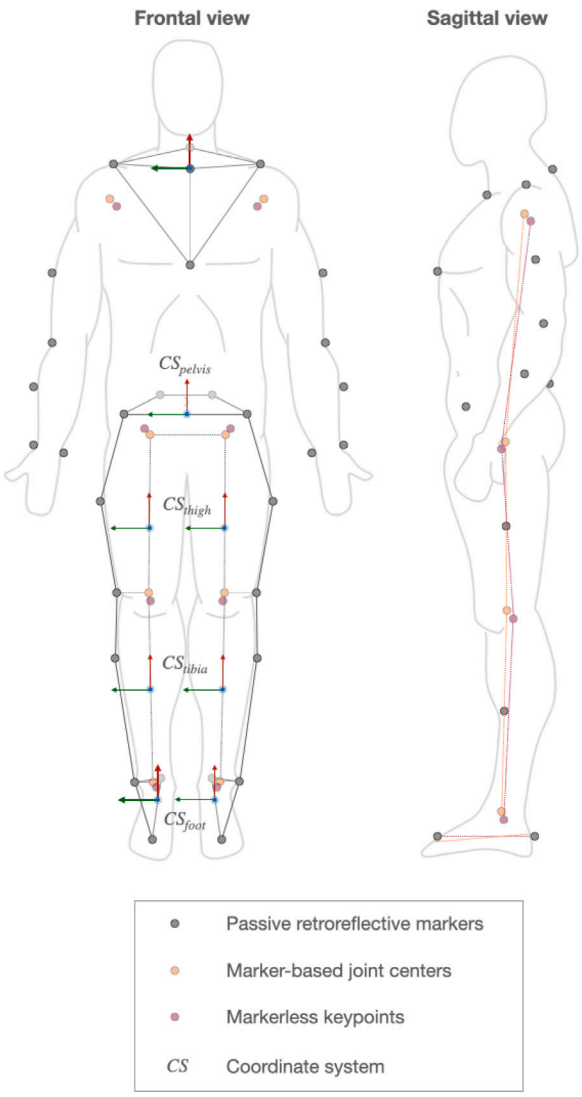

Fig. 3. Physical markers (gray circles) placed on the subjects, the marker-based joint centers (yellow circles) and the markerless keypoints (pink circles), obtained with the human pose estimation framework. The frontal view also shows the local coordinate systems (CS) for each segment; while the sagittal view shows the lines approximating segments to calculate sagittal 1D-kinematics.

2.2.3. Systems synchronization

To ensure data synchronization and be able to compare output from both MB and ML motion capture systems, an electrical signal connected to a LED lamp served as trigger as in [21]. When the LED was turned on, the electrical signal was fed to the Vicon analog box to be recorded, whilst the red light was recorded by the RGB camera.

2.2.4. Motion task description

Participants were instructed to walk self-paced on a straight line back and forth along a 6 m pathway within the capture volume of the MB-system for 10 times (Fig. 2). In particular: they walked from A to B; turned around the cone B; walked back from B to A; turned around the cone A; and started again, repeating such tasks until they travelled 20 straight full tracks (i.e., 10 per exposed side to the camera). The choice of a so relatively large number of straight tracks to be collected is justified by the fact that single-camera ML systems tend to fail reconstructing the position of keypoints when human segments visibility is occluded, for example due to overlapping of lower limbs while walking [17]. Data from the straight tracks, concurrently gathered with the MB and ML systems, were used to test the effect of number of considered strides on the precision of joint kinematics estimates.

2.3. Processing of human joint kinematics

2.3.1. Marker-based motion capture

Labelling of marker trajectories, gap filling, smoothing (Woltring spline routine, size 30 [37]), and manual gait cycle events detection (i.e., instants of foot-strike and foot-off) were conducted within Vicon Nexus (v2.14, Vicon, Oxford, UK). Foot-strike and foot-off events are defined as the instants when the foot first touches the ground while walking (with the heel in normal walking) and leaves the ground (with the toes in normal walking), respectively [38]. Gait cycle events were defined by visual inspection of trajectories of heel and toe markers. Two additional events were saved to ease the post-processing and trim each walking trial into the 20 straight tracks: the IN event (i.e, the first foot-strike of the straight walking with full visible lower limb joints, with tolerance of 10 samples - 0.08 s) and the OUT event (i.e, the last foot-strike with full visible lower limb joints, with tolerance of 10 samples - 0.08 s).

The local coordinate systems for each segment were defined based on markers instantaneous position and according to the PiG model [32], and used to: (i) estimate the 3D coordinates of shoulder (SJC), hip (HJC), knee (KJC) and ankle (AJC) joint centres; and (ii) compute the 3D hip, knee and ankle joint kinematics [4,32].

Data were then exported as MAT-files using custom-made scripts based on the integrated MATLAB SDK provided with Vicon Nexus (v2.14, Vicon, Oxford, UK). Only sagittal kinematics were then retained for the analysis.

2.3.2. Markerless motion capture

Images collected with the RGB-D sensor were fed to the framework OpenPose [33] to infer the 2D coordinate of the shoulder, hip, knee and ankle joint centers at each collected frame. 2D coordinates were then augmented to the 3D space using data from the depth sensor.

Starting from the extrapolated keypoints $\{JC_i\}$, with $i = 1, \ldots, N$, the sagittal kinematics of the jth joint, with center in JC_j, was estimated as the angle between the intersecting lines connecting JC_j with the adjacent human joint centers JC_{j-1} and JC_{j+1} on the same side. As an example, the knee flexion/extension kinematics would be computed as the angle between the line connecting the hip (A) and the knee (B) joint center (i.e., the proximal body segments – $prox$), and the line connecting the knee to the ankle (C) joint center (i.e., the distal body segment – $dist$). See Fig. 3. The angle insisting on point B is defined obtained via the following steps:

- build the segments' unit vectors $\hat{ab}$ and $\hat{bc}$,

$$\begin{aligned} \hat{ab} &= \|B - A\|, \quad \text{unit vector connecting } A \text{ to } B \\ \hat{bc} &= \|C - B\|, \quad \text{unit vector connecting } B \text{ to } C \end{aligned} \tag{1}$$

- build the local Coordinate System solid with the proximal segment and centered on the joint B, with orientation matrix ${}^{prox}\mathbf{R}_{dist}$:

$$\begin{aligned} \hat{j} &= \hat{ab} \times \hat{bc} \\ \hat{i} &= \hat{j} \times \hat{k} \\ \hat{k} &= -\hat{ab} \end{aligned} \tag{2}$$

$${}^{prox}\mathbf{R}_{dist} = \begin{bmatrix} | & | & | \\ \hat{i} & \hat{j} & \hat{k} \\ | & | & | \end{bmatrix} \tag{3}$$

- the 1D angle, approximating the sagittal joint kinematics, is then equal to the orientation angle of $\hat{bc}$ on the (i,k)-plane of this Cooridinate System:

$$\begin{aligned} {}^{prox}\mathbf{R}_{dist} \cdot \hat{bc} &= \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ \theta &= \text{atan2}(z, x) \end{aligned} \tag{4}$$

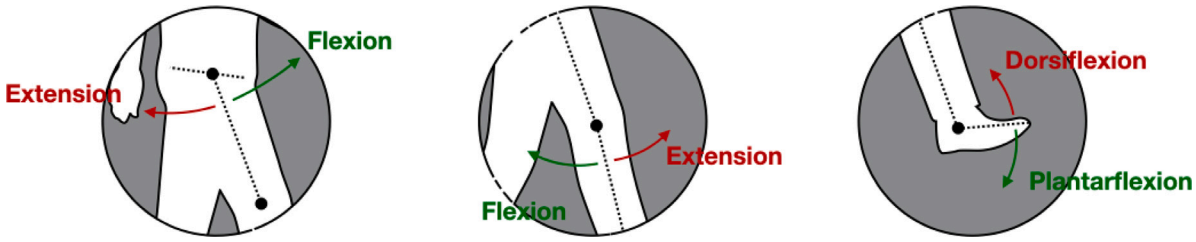

Fig. 4. Hip, knee and ankle sagittal joint kinematics (from left to right).

Table 1
Description of sagittal joint kinematics calculated according to: the Plug-in-Gait 3D model [32] (MB-PiG); the 1D geometrical approach based on the PiG-based joint centers (MB-JC); and the 1D geometrical approach based on the keypoints extrapolated with the human pose estimation framework, approximating the human joint centers (ML-JC).

Kinematics variable	MB-PiG	MB-JC	ML-JC
Hip flex/extension	sagittal rotation from the Cardan decomposition of the rotation matrix between CS_{pelvis} and CS_{thigh}	between the lines joining ${}^{MB}(HJC - SJC)$, and ${}^{MB}(HJC - KJC)$	between the lines joining ${}^{ML}(HJC - SJC)$, and ${}^{ML}(HJC - KJC)$
Knee flex/extension	sagittal rotation from the Cardan decomposition of the rotation matrix between CS_{thigh} and CS_{tibia}	between the lines joining ${}^{MB}(KJC - HJC)$, and ${}^{MB}(KJC - AJC)$	between the lines joining ${}^{ML}(KJC - HJC)$, and ${}^{ML}(KJC - AJC)$
Ankle plantar/dorsiflexion	sagittal rotation from the Cardan decomposition of the rotation matrix between CS_{tibia} and CS_{foot}	between the line ${}^{MB}(AJC - KJC)$ and the line joining ${}^{MB}HEEL$ and ${}^{MB}TOE$ markers	between the line ${}^{ML}(AJC - KJC)$ and the line ${}^{ML}(HEEL - FF)$, with FF being the midpoint between first and fifth foot fingers

It is worth underlying that the joint kinematics obtained from the ML system is only defined on the camera image plane (i.e., 1D angles), which can be assumed as coincident to the sagittal plane when the subject walks on a line parallel to that plane, whereas marker-based kinematics is generally three-dimensional and calculated by modelling the human joints as mechanical joints (i.e., hinge, spherical, Cardan joints, etc. — see Table 1, Figs. 3 and 4 for details) [4]. This intrinsic difference in calculating the joint angles through MB and ML systems adds a bias when comparing sagittal joint kinematics obtained with these systems. Fig. 5 shows the hip and knee flexion/extension and ankle plantar/dorsiflexion angles calculated (i) according to the PiG biomechanical model, (ii) from the marker-based joint centers, and (iii) from the HPE keypoints.

Each kinematic signal extrapolated from the ML keypoints was processed through a resampling and a filtering phase. In detail, a 3rd order Savitzky–Golay filter (i.e., a moving window-based smoothing) was first used to smooth the ML-based kinematic time-histories. Subsequently, time-histories were resampled to a fixed frequency equal to 60 Hz. This step is essential to address any potential fluctuations in camera sampling rates. Eventually, kinematics were low-passed filter with a 4th-order zero-lag Butterworth filter with a cut-off equal to 6 Hz, to remove high-frequency noise and outliers.

2.4. Kinematic features extraction

For the Reader convenience, a tag corresponding to the computation approach is used as left superscript for each kinematic variable $\alpha_{J_j}(t)$, with $J_j \in \{hip,\ knee,\ ankle\}$. As an example, the hip flexion/extension angle will be addressed as: ${}^{MB-PiG}\alpha_{hip}(t)$, when computed as Euler/Cardan angle between two adjacent segments, as defined in the PiG biomechanical model; and ${}^{MB-JC}\alpha_{hip}(t)$ or ${}^{ML-JC}\alpha_{hip}(t)$, when computed as the 1D angle between two lines connecting the hip joint center to the adjacent (in this case, shoulder and knee) joint centers (JC), defined with the MB and ML systems respectively.

The portions of straight walking of each trial ($\{SW_{k,s}\}$, with $k = 1,\ \ldots,\ 10$ and $s = l,\ r$ for the left and right side exposed to the camera while walking) gathered from each participant were extracted considering the IN and OUT events. For each $\{SW_{k,s}\}$, kinematics of the lower limb not directly exposed to the camera was also computed and analysed.

Within each $SW_{k,s}$, marker time-histories were partitioned according to the 10 sectors defined with the floor markers and considering a sliding w-width window consisting of 3 sectors at a time (i.e., $w = 1.5$ m) – see Fig. 2c. The w was chosen to consider the joints kinematics of an entire gait cycle, which is normally 1 m long for a population of healthy adults [39].

For each positioning of the sliding window ($\{W_t\}$ = {I-II-III, II-III-IV, III-IV-V, …, VIII-IX-X}), the following summary metrics [38,40] were computed and stored for further analysis (Fig. 5):

- the articular range of motion $\mathrm{ROM}_{J_j}\left|\left(\alpha_{J_i}(t)\right)\right.$, defined as max $\left(\alpha_{J_j}(t)\right)$ - min$\left(\alpha_{J_i}(t)\right)$;
- the maximum flexion (H_1) and maximum extension (H_2) of the hip over the gait cycle;
- the maximum knee flexion in mid-stance (K_1);
- the maximum knee extension in terminal stance (K_2);
- the maximum knee flexion in swing (K_3);
- the maximum ankle dorsiflexion in terminal stance (A_1);
- the maximum ankle plantar flexion (A_2) and maximum dorsiflexion (A_3) in swing.

MB-based data (both MB-PiG and MB-JC) were compared to those obtained with the ML system. The average values of each quantity among all the $SW_{k,s}$ and all the gait cycles from the MB measures served as ground truth for each participant. Average values and relevant confidence intervals of each quantity obtained from the ML measures will rather serve as reference values for clinicians who will be assessing individuals' ML-JC kinematics.

In order to test whether it exists a number $\bar{k} \leq 10$ of straight walking tracks to reach a plateau in the measurement precision of the considered summary metrics, average and standard deviation of each quantity were calculated in each W_t, for each participant, with $k = 1,\ \ldots,\ \bar{k}$ and for both sides (i.e., left and right).

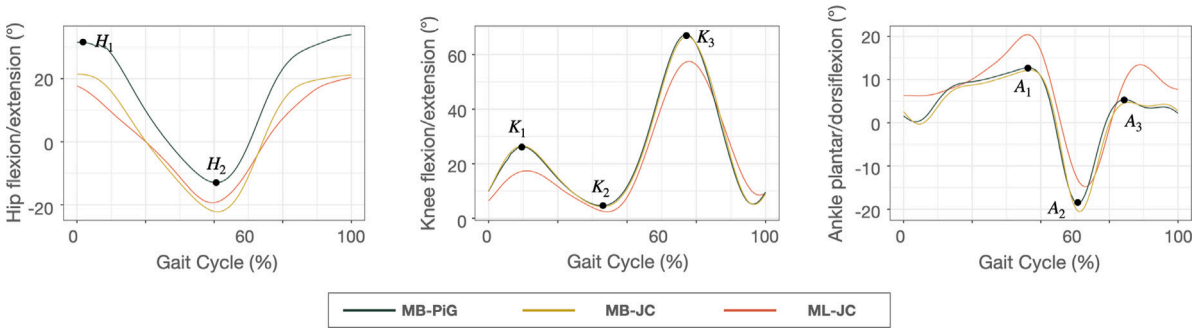

Fig. 5. Hip, knee and ankle joint angles on the sagittal plane over the gait cycle percentage (i.e., between subsequent foot-strikes events of the same foot) calculated according to the relevant biomechanical model (Plug-in-Gait - MB-PiG, dark green line), starting from the marker-based joint centers (MB-JC, dark yellow line), and from the human pose estimation keypoints (ML-JC, dark orange line).

Guidelines to use the ML system will therefore include the requirement for the subject to walk $\bar{k}$ times in front of the camera with one side of his/her body exposed.

Data processing was performed with custom made Python programs (Python 3, Python Software Foundation, 9450 SW Gemini Dr, ECM# 90772, Beaverton, OR 97008, USA).

2.5. Data analysis

2.5.1. Statistical tests on joint kinematics summary metrics

Having fixed the $\bar{k}$ of needed straight walking, differences among MB-PiG, MB-JC and ML-JC measures for each W_t were tested either with a two-ways repeated measures ANOVA or two Friedman tests (one to test the effect of moving among W_t, while fixing the method to calculate the kinematics, and one to test the effect of changing method to calculate kinematics within a W_t). Significance level was set to 0.05. Pairwise comparisons were then performed via paired two-tails t-test, or paired two-tails Wilcoxon rank-sum test, considering a Bonferroni correction: i.e., the significance level was set to 0.017 when performing pairwise comparisons among methods to compute kinematics (3 multiple comparison within the single W_t); and to 0.0018 when comparing data gathered among different W_t (28 multiple comparison within calculation method). The choice between the parametric and non-parametric tests, both for the omnibus tests and the pairwise comparisons, was based on data normality checked with a Shapiro–Wilk test ($p > 0.05$) [41]. Statistical tests were run in R-Studio v2023.03.0+386 [42].

2.5.2. Whole-cycle kinematics comparison

Outputs obtained from a MB approach (i.e., joint kinematics time-histories) strongly depend on the adopted model, which consists of the number and location of markers to be placed on the subject to be studied and the biomechanics model used to solve the inverse kinematics [4,43,44]. Thus, looking only at absolute differences among angles obtained with MB and ML systems would not be completely meaningful [45]. Moreover, the most relevant quality of gait analysis outputs is for them to be repeatable (i.e., measurement precision associated with the same operator performing the same procedure on the same group of subjects that in gait analysis quantifies the within- and between-subject variability) and reproducible (i.e., the measurement precision associated with different operators performing the same procedure on the same group of subjects that quantifies the between-operator variability of the data) [46].

With these perspectives, having chosen and fixed $\bar{k}$, sagittal kinematics were compared, looking at absolute agreement among computation methods with the Linear Fit Method (LFM) [47]. The LFM calculates the scaling factor (a_1), the weighted averaged offset (a_0), and the trueness (R^2) of the linear regression model between a set of curves and a reference curve. The coefficients a_1 tends to 1 for perfect matching of motion amplitude, and a_0 to 0 when in absence of any offset among the comparing curves [47]. The ideal value for R^2 is 1, and a decrease of this parameter value shows the presence of a time shift between the compared curves. Care should be paid to interpreting the remaining coefficients and absolute differences [45]. LFM was fed with the averaged stride curve obtained for each participant and each joint kinematics (time-normalized over the percentage of the gait cycle) obtained from ML and MB systems, with the latter taken as a reference. To test whether the precision of ML-based joint kinematics increases with the video recording being taken in the centre of the capture volume, the LFM analysis was repeated considering both the whole straight walking track $SW_{\bar{k},s}$ and what was considered as the "optimal view zone" W_t, i.e. from sector III to VIII in Fig. 2c. The same analyses were run both comparing the MB-PiG with ML-JC, and the MB-JC and ML-JC kinematics.

As a further test, joint kinematics averaged within-participants were compared with the grand mean between-participants via LFM, both for the whole straight walking and the "optimal view zone". This would allow testing the usability of such ML approach in clinical practice, as patient data are normally superimposed to normative bands. Indeed, normative bands thickness built on a healthy population should not hinder possible kinematics alterations due to pathologies. The test was run for both the exposed and occluded side of the body to the ML camera.

2.5.3. Bland–Altman analysis on joint kinematics summary metrics

To complement the averaged offset (LFM-a_0) information and better appreciate the bias induced in joint kinematics by the different calculation methods at specific instants of the gait cycle, a Bland–Altman analysis [48,49] was run on the summary metrics given in Fig. 5 and paragraph 2.4, having considered the optimal view zone. The Bland–Altman analysis returns the interchangeability (i.e., the bias and limit of agreement — normally set to 95%) between two methods used to measure or estimate a quantity. The Bland–Altman analysis will favour clinicians' understanding of switching from the more familiar MB to ML measures. A bias between the two methods should be read as the intrinsic geometrical difference in calculating each variable, as in paragraph 2.3.2 and Fig. 3.

3. Results

From visual inspection of results obtained for each summary metric of each joint (see Supplemental Materials), no clear trend of decreasing or increasing variations (in terms of standard deviation and closeness to the golden standard — the MB-based measures) was obtained. The number $\bar{k}$ of straight walking tracks to be travelled in front of the ML camera was then set equal to 5 as generally done for gait analysis performed with MB systems [29]. The results presented hereinafter are obtained having fixed $\bar{k} = 5$.

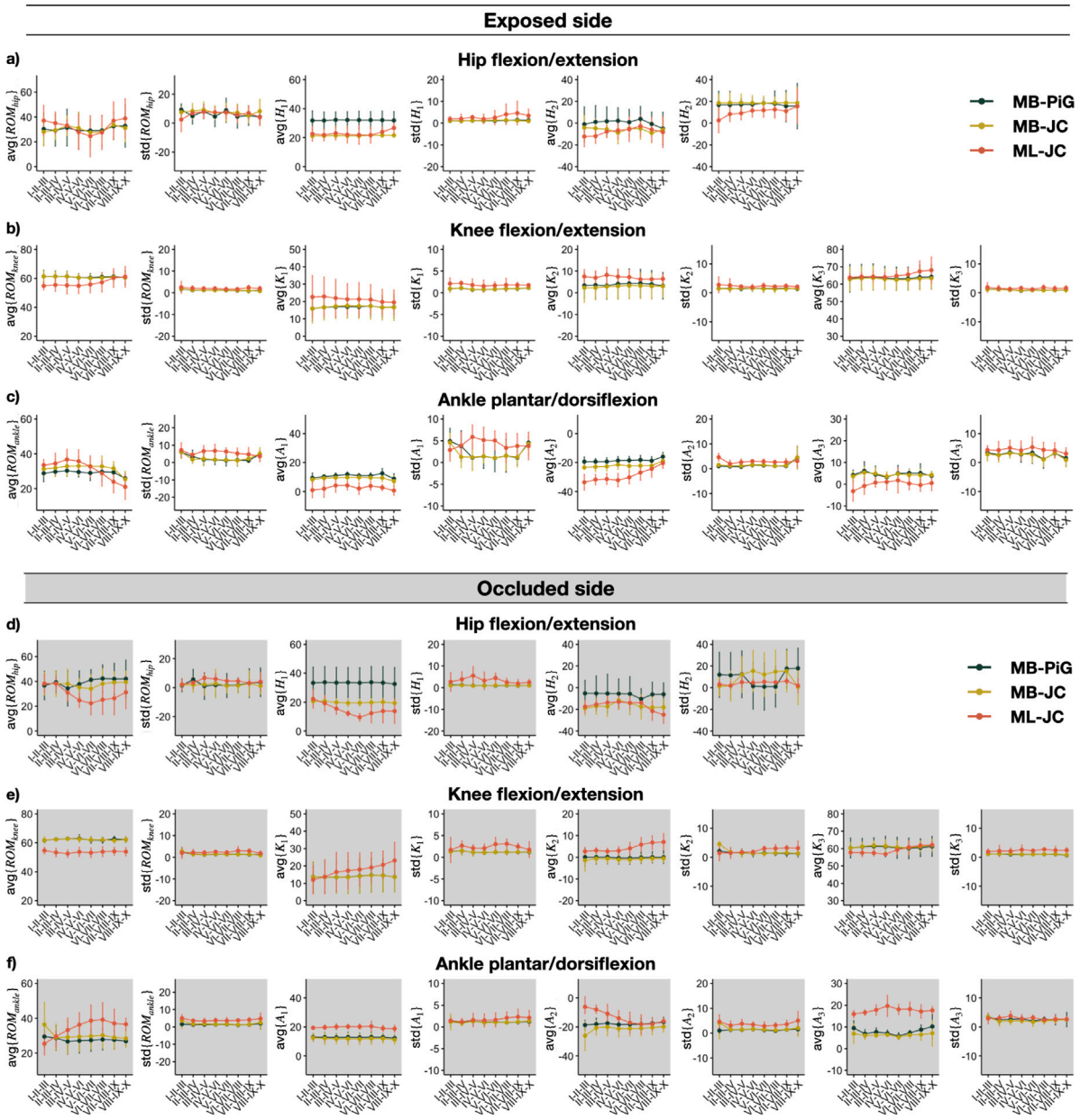

Fig. 6. Summary metrics calculated for hip, knee and ankle joints computed according to MB-PiG (dark green), MB-JC (dark yellow) and ML-JC (dark orange) methods, for both the exposed (white background - panels a), b) and c)) and occluded side (gray background - panels d), e) and f)) of the body concerning the markerless (ML) camera. Each summary metric is presented as averages over strides within-participants (avg{·}), and standard deviations over strides within-participants (std{·}), both are depicted as the median and inter-quartile range over participants. The *x*-axis reports the sectors used to obtain the results in Roman numerals. From top to bottom (from panel a) to panel c) for the exposed side; and from panel d) to panel f for the occluded side), the summary metrics are grouped by human joint kinematics: hip flexion-extension (panels a and d), knee flexion-extension (panel b and e), and ankle plantar/dorsiflexion (panels c and f).

Fig. 6 shows the results obtained for all the summary metrics, computed for each joint and according to all the methods (i.e., MB-PiG, MB-JC and ML-JC), for both the exposed and occluded side of the body concerning the ML camera. For each summary metric, Fig. 6 provides the Reader with (i) values averaged over strides within-participants ($avg\{\cdot\}$), and (ii) standard deviations over strides within-participants ($std\{\cdot\}$), both presented as median and inter-quartile range over participants. Looking at these results, both for the within-participants averaged values and the within-participants standard deviations for all the summary metrics, it seems that a trend towards a lower precision (i.e., larger dispersion) was obtained when looking at the boundary of the capture volume (i.e., I–II–III, and VI–VII–VIII and VIII–IX–X sectors).

The Shapiro–Wilk tests revealed a non-normal distribution for some of the summary metrics ($p > 0.05$, see Supplemental Materials for details). Non-parametric tests were then considered for all the comparisons. Statistical analysis results, including normality tests, are extensively presented in the Supplemental Materials, while a few specific details are provided below.

The Friedman tests found differences for the majority of the comparisons that were run when changing the observation block (the tth position of W_t) and considering the methods to calculate the human joint kinematics as the main effect ($p < 0.05$). Differences were still found after the Bonferroni correction to operate the pairwise comparison ($p_{adj} < 0.05$). When considering the computation method and the effect of changing the observation block, the Friedman tests

Table 2
Comparison of kinematics obtained from the considered marker-based biomechanical model (i.e., Plug-in-Gait) and from the markerless-based measures returning Linear Fit Method (LFM) coefficients: the scaling factor (a_1), the weighted averaged offset (a_0), and the trueness (R^2) of the linear regression model, considering the whole straight walking track (between sectors I to X) as testing curves and the "optimal view zone" (between sectors III to VIII) as reference curve. LFM coefficients are computed for each participant and reported as mean and standard deviation.

MB-PiG vs ML-JC						
	Whole straight walking (I-X)			Optimal view zone (III-VIII)		
	a_1	a_0 (°)	R^2	a_1	a_0 (°)	R^2
	Exposed side					
Hip flex/extension	0.77 ± 0.07	−3.7 ± 4.0	0.96 ± 0.03	0.74 ± 0.07	−4.3 ± 4.9	0.95 ± 0.03
Knee flex/extension	0.90 ± 0.07	4.9 ± 3.6	0.94 ± 0.03	0.89 ± 0.08	4.0 ± 4.0	0.95 ± 0.03
Ankle plantar/dorsiflexion	1.00 ± 0.14	−3.1 ± 7.5	0.84 ± 0.09	0.97 ± 0.22	−2.5 ± 6.2	0.80 ± 0.11
	Occluded side					
Hip flex/extension	0.71 ± 0.07	−3.4 ± 3.9	0.96 ± 0.02	0.64 ± 0.09	−2.6 ± 4.2	0.92 ± 0.06
Knee flex/extension	0.90 ± 0.06	4.9 ± 3.0	0.95 ± 0.02	0.89 ± 0.07	5.5 ± 3.4	0.95 ± 0.03
Ankle plantar/dorsiflexion	1.01 ± 0.20	−4.3 ± 9.5	0.80 ± 0.17	1.03 ± 0.17	−3.8 ± 7.6	0.82 ± 0.10

Table 3
Comparison of kinematics obtained from the joint centers obtained with the marker-based biomechanical model (i.e., Plug-in-Gait) and from the markerless-based measures returning Linear Fit Method (LFM) coefficients: the scaling factor (a_1), the weighted averaged offset (a_0), and the trueness (R^2) of the linear regression model, considering the whole straight walking track (between sectors I to X) as testing curves and the "optimal view zone" (between sectors III to VIII) as reference curve. LFM coefficients are computed for each participant and reported as mean and standard deviation.

MB-JC vs ML-JC						
	Whole straight walking (I-X)			Optimal view zone (III-VIII)		
	a_1	a_0 (°)	R^2	a_1	a_0 (°)	R^2
	Exposed side					
Hip flex/extension	0.82 ± 0.08	1.5 ± 3.5	0.96 ± 0.03	0.79 ± 0.08	0.6 ± 4.7	0.95 ± 0.03
Knee flex/extension	0.90 ± 0.07	5.0 ± 3.2	0.95 ± 0.03	0.89 ± 0.08	4.1 ± 3.7	0.95 ± 0.03
Ankle plantar/dorsiflexion	0.98 ± 0.13	−1.3 ± 6.0	0.85 ± 0.08	0.94 ± 0.21	−1.1 ± 6.1	0.80 ± 0.13
	Occluded side					
Hip flex/extension	0.78 ± 0.07	3.0 ± 2.8	0.96 ± 0.02	0.70 ± 0.09	3.2 ± 3.1	0.93 ± 0.05
Knee flex/extension	0.90 ± 0.06	4.9 ± 2.8	0.95 ± 0.02	0.89 ± 0.07	5.5 ± 3.3	0.95 ± 0.03
Ankle plantar/dorsiflexion	0.94 ± 0.12	−1.9 ± 7.5	0.83 ± 0.10	0.97 ± 0.16	−2.6 ± 7.5	0.81 ± 0.10

highlighted some differences ($p < 0.041$). However, having run 27 pairwise comparisons ($p_{limit} = 0.0018$), no differences survived to the Bonferroni correction ($p_{adj} > 0.437$). A trend towards significance is observed for some summary metrics, mainly associated with ankle kinematics, and supported by not-adjusted p-values equal to 0.016.

Tables 2 and 3 show the results of comparisons of the ML-JC kinematics with those obtained with the classical MB-PiG and with the MB-JC approaches, respectively. In general, excellent correlation was obtained for hip and knee flexion-extension angle in both comparisons and considering both the exposed and the occluded side of the body to the ML camera ($0.92 < R^2 < 0.96$), with the lowest R^2 obtained for the hip flexion-extension angle on the occluded side (MB-PiG vs ML-JC: $R^2 = 0.92 \pm 0.06$; MB-JC vs ML-JC: $R^2 = 0.93 \pm 0.05$). Slightly lower values of correlation were obtained for the ankle plantar-dorsiflexion. The a_1 values get closer to 1 when going from comparing ML-JC with MB-PiG to MB-JC in both Tables 2 and 3. As for the a_0, values do not significantly change when looking at knee flexion-extension for both exposed and occluded sides of the body, between MB-PiG and MB-JC comparisons with the ML-JC kinematics. Differently, averaged values and standard deviations for a_0 decreased from MB-PiG to MB-JC compared to ML-JC based hip and ankle kinematics, both for the exposed and the occluded sides of the body.

Table 4 reports the comparison between kinematics obtained for each participant to the grand mean among all the participants. The average correlation among participants ranged between 0.92 and 0.94 for hip and knee flexion-extension, with an averaged shift (i.e., a_0) close to 0 (with a standard deviation $\leq 5.6°$) for hip and knee flexion-extension. When considering the ankle plantar-dorsiflexion, correlation lowers in the range from 0.76 to 0.81, with an averaged shift in the kinematics among participants reaching $0.8° \pm 11.2°$ for the side of the body occluded to the camera objective.

The Bland–Altman analysis returned complementary information to the averaged offset (LFM-a_0), whose results are provided in Table 5 and graphically given in the Supplemental Materials. The first column of Table 5 shows the effect of comparing the joint kinematics summary metrics gathered MB-PiG and MB-JC measures, quantifying the effect of considering 3D modelling rather than 2D modelling for the human joints. The larger bias (in the range [0.2°, 10.9°]) was obtained for the hip flexion/extension, with a width for the limit of agreement at 95% of confidence in the range between 13.0° and 33°. The smallest effect of changing between MB-PiG and MB-JC was obtained for the knee joint, with a bias (in absolute values) ranging between 0.1° and 0.8° and width of the limit of agreement always lower than 4°. The bias introduced in the ankle kinematics ranged between 0.8° and 2.7°, with a limit of agreement width between 2.7° and 8.4°. The MB-JC and ML-JC comparison highlighted the largest bias introduced to the ankle joint kinematics calculation: between 1.2° and 11.8°, with a width of limit of agreement between 4.6° and 15.6°. The smallest bias was obtained for the hip joint: between 0.2° and 7.3°, with a width of the limit of agreement between 11.0° and 39.0°. The bias on the knee kinematics was instead between 1.2° and 8.4°, with a width of limit of agreement between 7.1° and 13.3°. The comparison between the 3D MB-PiG and the 2D ML-JC highlighted the largest bias introduced to the ankle joint kinematics calculation: between 2.5° and 10.6°, with a width of limit of agreement between 6.7° and 19.0°. The smallest bias was obtained for the knee joint: between 1.0° and 9.0°, with a width of the limit of agreement between 8.6° and 17.1°. The bias on the hip kinematics was between 1.5° and 17.3°, with a width of the limit of agreement between 17.2° and 53.0°.

Table 4
Comparison of kinematics obtained from the joint centres obtained from the markerless-based measures for each participant with the grand mean over all the participants, returning Linear Fit Method (LFM) coefficients: the scaling factor (a_1), the weighted averaged offset (a_0), and the trueness (R^2) of the linear regression model, considering the whole straight walking track (between sectors I to X) as testing curves and the "optimal view zone" (between sectors III to VIII) as reference curve. LFM coefficients are computed for each participant and reported as mean and standard deviation.

ML-JC (single sbj) vs ML-JC (grand mean)						
	Whole straight walking (I-X)			Optimal view zone (III-VIII)		
	a_1	a_0 (°)	R^2	a_1	a_0 (°)	R^2
	Exposed side					
Hip flex/extension	1.03 ± 0.13	0.9 ± 5.2	0.94 ± 0.04	0.99 ± 0.12	0.2 ± 5.6	0.93 ± 0.06
Knee flex/extension	1.00 ± 0.12	0.9 ± 4.4	0.92 ± 0.09	1.00 ± 0.13	0.0 ± 4.5	0.92 ± 0.10
Ankle plantar/dorsiflexion	1.07 ± 0.24	−0.8 ± 9.1	0.79 ± 0.17	1.04 ± 0.31	−0.3 ± 9.0	0.76 ± 0.17
	Occluded side					
Hip flex/extension	1.12 ± 0.20	−0.4 ± 2.0	0.94 ± 0.05	1.01 ± 0.21	0.2 ± 2.7	0.91 ± 0.06
Knee flex/extension	1.03 ± 0.11	−1.1 ± 3.6	0.94 ± 0.06	1.02 ± 0.11	−0.5 ± 3.9	0.94 ± 0.07
Ankle plantar/dorsiflexion	1.01 ± 0.26	0.8 ± 11.2	0.83 ± 0.21	1.05 ± 0.29	0.5 ± 11.5	0.81 ± 0.21

Table 5
Comparison of sagittal hip, knee and ankle joints kinematics obtained from the considered marker-based biomechanical model (i.e., Plug-in-Gait - MB-PiG), the marker-based joint centers (MB-JC) and from the markerless-based measures returning Bland–Altman mean differences and 95% limit of agreement (in square brackets) for each summary metric for both the exposed (white background cells) and occluded side (gray background cells) of the body with respect to the markerless (ML) camera.

			MB-PiG vs MB-JC	MB-JC vs ML-JC	MB-PiG vs ML-JC
Hip flex/extension	ROM_{hip}	Exposed	0.2 [−16.0, 17.0]	−1.8 [−21.0, 18.0]	−1.5 [−26.0, 23.0]
		Occluded	2.5 [−4.0, 9.1]	7.3 [−5.4, 20.0]	9.8 [−3.8, 24.0]
	H_1	Exposed	8.6 [−2.4, 19.0]	−0.2 [−7.4, 7.0]	8.4 [−0.2, 17.0]
		Occluded	10.9 [−2.7, 24.0]	6.4 [1.1, 12]	17.3 [5.3, 29.0]
	H_2	Exposed	7.6 [−9.0, 24.0]	1.7 [−18.0, 21.0]	9.3 [−17.0, 36.0]
		Occluded	7.7 [−7.3, 23.0]	−0.4 [−13.0, 12.0]	7.3 [−12.0, 26.0]
Knee flex/extension	ROM_{knee}	Exposed	0.1 [−0.5, 0.6]	3.7 [−1.2, 8.6]	3.8 [−1.1, 8.8]
		Occluded	0.6 [−0.7, 2.0]	8.4 [2.6, 14.0]	9.0 [4.0, 14.0]
	K_1	Exposed	−0.3 [−2.1, 1.4]	−1.2 [−5.3, 2.8]	−1.6 [−6.4, 3.2]
		Occluded	−0.2 [−1.6, 1.2]	1.2 [−7.2, 9.6]	1.0 [−7.0, 9.1]
	K_2	Exposed	0.3 [−0.9, 1.5]	−6.1 [−13.0, 0.3]	−5.9 [−13.0, 1.6]
		Occluded	0.2 [−0.7, 1.2]	−4.7 [−11.0, 1.5]	−4.5 [−11.0, 2.2]
	K_3	Exposed	0.3 [−0.9, 1.6]	−2.3 [−5.9, 1.2]	−2.0 [−6.3, 2.3]
		Occluded	0.8 [−1.2, 2.8]	3.7 [−5.0, 12.0]	−4.5 [−4.1, 13.0]
Ankle plantar/dorsiflexion	ROM_{ankle}	Exposed	−1.3 [−5.5, 2.9]	−1.2 [−9.0, 6.6]	−2.5 [−12.0, 7.0]
		Occluded	−1.1 [−3.9, 1.6]	−3.2 [−8.1, 1.7]	−4.3 [−11.0, 2.0]
	A_1	Exposed	1.4 [0.1, 2.7]	6.2 [2.0, 10.0]	7.5 [3.6, 11.0]
		Occluded	1.3 [0.2, 2.5]	−9.5 [−14.0, −5.5]	−8.2 [−12.0, −4.8]
	A_2	Exposed	2.7 [−0.7, 6.2]	7.4 [1.0, 14.0]	10.1 [1.4, 19.0]
		Occluded	2.5 [0.5, 4.5]	−6.3 [−11.0, −1.6]	−3.8 [−10.0, 2.4]
	A_3	Exposed	0.8 [0.9, 2.4]	2.8 [−2.9, 8.5]	3.6 [−2.6, 9.7]
		Occluded	1.2 [−0.2, 2.7]	−11.8 [−14.0, −9.4]	−10.6 [−14.0, −7.3]

4. Discussion

This study emphasizes the potential of single-camera markerless systems in telemedicine applications to assess human joint kinematics during overground walking. The relevance of motion analysis and its potential in telemedicine applications has been well established [4–7,17,18]. The integration of markerless motion capture systems into telemedicine platforms has the potential to remodel remote evaluation and rehabilitation effectiveness, enabling personalized interventions while reducing the relevant healthcare costs and increasing accessibility and adherence to therapies for individuals in need [17,18,50]. Questions have been posed on future healthcare to fully embrace widespread use of IoT and telemedicine for remote patient monitoring and delivery of interventions, also associated with security of patients' data being shared and stored in clouds or similar platforms [51]. Other questions arise from a lack of clarity in definitions and how technologies should be used for that purpose, including among the scientific community [50]. Besides the concerns about taxonomy and the establishment of definitions, the usability of systems has also come under analysis [50].

Although many user-friendly devices have been proposed in the past years, their application in telemedicine may have been hampered by scepticism on the actual usability of the measures in clinics [25]. The present study aimed to fill this gap by providing end-users with instructions to use a single-camera ML system to provide clinicians with reliable motion data for continuous monitoring of walking, particularly in a telemedicine context. Towards this aim, kinematics gathered from participants walking self-paced overground and estimated using a markerless (ML) systems were compared to those obtained from marker-based (MB) systems by varying the number of considered strides and the positioning of the targeted subject within the capture volume. Namely, these research aims can be translated into answering the following questions.

***RQ1**: How many times should a user walk back and forth in front of the ML system to collect sufficient data to obtain reliable kinematics?*

Inspection of the results derived from the summary metrics for each joint kinematics, which were calculated by incrementally adding walking tracks for analysis (refer to the Supplemental Materials), no discernible trend of decreasing or increasing standard deviation emerged. Similarly, there was no apparent convergence to the ground truth values, represented by the measures derived from model-based (MB) approaches. In line with established clinical practices using MB systems, to capture the typical intra-subject variability [29], we opted to consider 5 straight walking tracks recorded in front of the ML camera for the analysis.

***RQ2**: Is there a part of the capture volume to be preferred to record motion data within?*

Results in Fig. 6, for some of the summary metrics, suggest a trend towards lower precision (i.e., larger dispersion) when examining the boundaries of the capture volume (i.e., I–II–III, VI–VII–VIII, and VIII–IX–X sectors). The omnibus statistical test revealed differences in kinematics computed from data gathered in different zones of the capture volume (Friedman's test: $p < 0.041$). However, having run 27 pairwise comparisons, Bonferroni's adjusted p-value was too strict, and no difference was eventually significant. A trend towards significance was observed for some summary metrics (not-adjusted $p = 0.016$), mainly associated with ML-JC kinematics obtained at the two boundary sides of the camera image plane. Discarding the portion of curves coming from data recorded at the boundaries of the camera image plane seems to be good practice. This last recommendation comes from the acknowledgement of possibly encountering two sources of error in these regions: (i) a reconstruction error, associated with the inability of the neural network to reconstruct the keypoints accurately, maybe due to occlusion of some part of the body [17]; and (ii) smoothing and filtering tail effects, which may introduce some distortion in the joint kinematics estimate. Moreover, it is worth underlying that the used ML system can be assumed to have no fish-eye distortion [52,53], which could not be the case when other ML cameras are considered, and further distortion could be introduced to kinematic data.

Overall, considering the participant's position within the capture volume appears appropriate, as well as discarding measurements obtained from the boundaries of the camera image plane.

***RQ3a**: Is ML kinematics reliable enough for clinical use? – **RQ3b**: Is it possible to consider data gathered from the side of the body partially occluded to the camera objective?*

As a first step to answer these questions, the model agreement was studied using the LFM [47].

It is important to emphasize that distinctions noted in human joint kinematics between ML and MB approaches cannot solely be attributed to technique-specific errors. They also stem from variations in defining bony segment poses and the conventions utilized for angle calculations, with 1D angles employed for MB-JC and ML-JC, and 3D Euler/Cardan angles for MB-PiG. It is worth explaining that differences observed in human joint kinematics between ML and MB approaches are not entirely cannot solely to be ascribed to technique-specific errors. They also stem from variations in defining bony segment poses and conventions used for angles calculation: i.e., 1D angles for MB-JC and ML-JC and 3D Euler/Cardan angles for MB-PiG (Fig. 5) [4]. These differences are also revealed by: (i) the statistical tests when comparing different computation methods for the kinematic summary metrics in an observation sector; and (ii) the Bland–Altman analysis, which highlighted biases for all comparisons between calculation methods and all summary metrics. Specifically, the Bland–Altman analysis facilitates the identification of biases induced by the different modelling approaches of human joints when comparing MB-PiG and MB-JC, even though both datasets originate from the same source, namely the MB system. This is evident in Fig. 3.

Although comparing solutions to different geometrical problems might sound peculiar, clinicians would benefit from these results by not expecting ML systems to provide the same results yielded by MB systems when considering (for example) the PiG biomechanical model. The excellent correlations obtained in Tables 2 and 3 still suggest that ML-based kinematics describe the same movement, yet differently defined, of MB-based approaches. The LFM-a_1 and a_0 coefficients, together with the Bland–Altman biases and limits of agreement, emphasize the above-described differences. Being the MB-JC approach geometrically defined as the ML-JC approach, differences observed in Table 3 and in the relevant column of Table 5 can be mainly ascribed to the differences in estimating the joint centre positions between MB (i.e., using regressions from skin-mounted markers) and ML (i.e., the keypoints estimated with the HPE framework) systems. In this case, the LFM-a_1 tend to 1 and LFM-a_0 decrease (Table 3) compared to values obtained from the MB-PiG vs. ML-JC kinematics testing (Table 2).

Table 4 reports the comparison results within the ML-JC kinematics and between-participants, providing the Reader with the variability to be accounted for when building normative bands to compare patient data. The repeatability of ML-JC based kinematics suggests that ML systems can be effectively used for clinical assessments, with values of LFM-a_1, a_0, and R^2 falling within the range reported in similar studies and for clinical practice [13,44]. It should be noted, however, that care should be adopted when analysing ankle angles and comparing patient data with reference data, as a_0 values suggest relatively large normative bands for this variable, as also reported in previous studies [13,44].

Biases and limit of agreement are generally larger for all summary metrics when moving from the MB measures to ML-JC kinematics. This agrees with the results obtained from the LFM, indicating that greater differences are introduced to the kinematics when moving from MB to ML approaches.

The findings of this study demonstrated the potential of markerless motion capture systems, specifically single-camera systems, in assessing human joint kinematics during overground walking. The markerless approach offers several advantages over traditional stereophotogrammetry systems, such as ease of use, cost-effectiveness, and the ability to be employed in clinical and unsupervised environments [11]. These systems eliminate the need for extensive subject preparation, expert operators, and controlled environments, allowing individuals to perform daily-life activities while being monitored [10]. Furthermore, markerless systems overcome limitations associated with marker placement, including soft-tissue artefact and landmark misrecognition, which are commonly encountered in traditional motion capture techniques [12–14].

The results of this study show that reliable motion data for continuous monitoring of walking can be obtained by simply filming someone walking using a single-camera markerless system and ensuring the camera is placed at a distance that allows: (i) good visibility of the whole body silhouette; and (ii) a meaningful length for the walking path to be covered (i.e. ≥2 m), to capture full strides for both sides of the body. It is worth noticing that images at the boundaries of the camera image plane should be discarded to avoid tail effects of filters. Following this procedure will make the obtained human joint kinematics reliable enough for clinical interpretation.

This study does have some limitations. It did not explore the effect on joint kinematics of changing cameras, camera characteristics (such as fish-eye distortion, different field-of-view, etc.) and settings (such as camera resolution). The adopted resolution was chosen to maximize the acquisition frame rate. Besides these aspects can be addressed as a limitation, the tested device was selected with: (i) non-fish-eye distortion, as these cameras are more commonly used; and (ii) the adopted resolution is lower than those available from everyday-use devices, such as smartphones, to consider the worst case scenario.

Different camera distances from the midline of the walking path were not tested, with this aspect potentially affecting the precision of the tested human pose estimation framework in reconstructing the keypoints coordinates and, consequently, the human joint kinematics. That said, the obtained results are valid for the specifically used ML system (the used camera, with its specific settings, complemented with OpenPose as a pose estimator) and not as a general result for every ML system. Nevertheless, the proposed assessment framework could be adopted for future evaluations of the validity of different ML systems used in clinical settings, and future tests on different ML systems are worth performing.

Last but not least, this research aimed to pave the way for using ML technologies in telemedicine. The present paper only considers a limited sample of healthy individuals. More research into the applicability of ML technology to analyse the motor skills of people with neurological and orthopaedic disorders in an unsupervised environment is worth performing. Further tests and analyses are also mandatory to mimic body parts occlusions, which might affect the data reliability and may be associated with the presence of walking aids, assistance from others or homes' furniture.

5. Conclusion

This research confirms the reliability of markerless kinematics in clinical applications, including telemedicine context. By following a few recommended procedures, healthcare professionals can acquire reliable motion data, making markerless systems valuable for remote people monitoring without requiring specialized operators. Future advancements should investigate the effect of placing the camera at different distances from the walking line and applying this technology to individuals with neurological or orthopaedic disorders, to evaluate their motor skills and estimate the pathology severity.

CRediT authorship contribution statement

Michele Boldo: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Roberto Di Marco:** Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Enrico Martini:** Conceptualization, Methodology, Writing – review & editing. **Mauro Nardon:** Writing – review & editing. **Matteo Bertucco:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Nicola Bombieri:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

None declared.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.108101.

References

[1] C.D. Mathers, G.A. Stevens, T. Boerma, R.A. White, M.I. Tobias, Causes of international increases in older age life expectancy, Lancet (ISSN: 1474547X) 385 (9967) (2015) 540–548, http://dx.doi.org/10.1016/S0140-6736(14)60569-9.

[2] M. Rubega, R. Di Marco, M. Zampini, E. Formaggio, E. Menegatti, P. Bonato, S. Masiero, A. Del Felice, Muscular and cortical activation during dynamic and static balance in the elderly: A scoping review, Aging Brain 1 (2021) 100013.

[3] M. van der Vlegel, J.A. Haagsma, A.J. Geraerds, L. de Munter, M.A. de Jongh, S. Polinder, Health care costs of injury in the older population: A prospective multicentre cohort study in the Netherlands, BMC Geriatrics (ISSN: 14712318) 20 (1) (2020) http://dx.doi.org/10.1186/s12877-020-01825-z.

[4] A. Cappozzo, U. Della Croce, A. Leardini, L. Chiari, Human movement analysis using stereophotogrammetry. Part 1: theoretical background, Gait Posture (ISSN: 0966-6362) 21 (2) (2005) 186–196, http://dx.doi.org/10.1016/j.gaitpost.2004.01.010, URL http://www.ncbi.nlm.nih.gov/pubmed/15639398.

[5] L. Chiari, U. Della Croce, A. Leardini, A. Cappozzo, Human movement analysis using stereophotogrammetry. Part 2: instrumental errors, Gait Posture (ISSN: 0966-6362) 21 (2) (2005) 197–211, http://dx.doi.org/10.1016/j.gaitpost.2004.04.004, URL http://www.ncbi.nlm.nih.gov/pubmed/15639399.

[6] A. Leardini, L. Chiari, U. Della Croce, A. Cappozzo, Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation, Gait Posture (ISSN: 0966-6362) 21 (2) (2005) 212–225, http://dx.doi.org/10.1016/j.gaitpost.2004.05.002, URL http://www.ncbi.nlm.nih.gov/pubmed/15639400.

[7] U. Della Croce, A. Leardini, L. Chiari, A. Cappozzo, Human movement analysis using stereophotogrammetry Part 4: Assessment of anatomical landmark misplacement and its effects on joint kinematics, Gait Posture (ISSN: 09666362) 21 (2) (2005) 226–237, http://dx.doi.org/10.1016/j.gaitpost.2004.05.003,

[8] V. Rajinikanth, S. Yassine, S.A. Bukhari, Hand-sketchs based parkinson's disease screening using lightweight deep-learning with two-fold training and fused optimal features, Int. J. Math., Stat., Comput. Sci. 2 (2024) 9–18.

[9] W.W. Lam, K.N. Fong, The application of markerless motion capture (MMC) technology in rehabilitation programs: A systematic review and meta-analysis, Virtual Reality (ISSN: 14349957) (2022) http://dx.doi.org/10.1007/s10055-022-00696-6.

[10] B. Scott, M. Seyres, F. Philp, E.K. Chadwick, D. Blana, Healthcare applications of single camera markerless motion capture: A scoping review, PeerJ (ISSN: 21678359) 10 (2022) http://dx.doi.org/10.7717/peerj.13517.

[11] M. Ota, H. Tateuchi, T. Hashiguchi, N. Ichihashi, Verification of validity of gait analysis systems during treadmill walking and running using human pose tracking algorithm, Gait Posture 85 (June 2020) (2021) 290–297, http://dx.doi.org/10.1016/j.gaitpost.2021.02.006.

[12] V. Camomilla, R. Dumas, A. Cappozzo, Human movement analysis: The soft tissue artefact issue, J. Biomech. (ISSN: 0021-9290) 62 (2017) 1–4, http://dx.doi.org/10.1016/j.jbiomech.2017.09.001, URL https://www.sciencedirect.com/science/article/pii/S0021929017304578, Human Movement Analysis: The Soft Tissue Artefact Issue.

[13] E. Scalona, R. Di Marco, E. Castelli, K. Desloovere, M. van Der Krogt, P. Cappa, S. Rossi, Inter-laboratory and inter-operator reproducibility in gait analysis measurements in pediatric subjects, Int. Biomech. 6 (1) (2019) 19–33, http://dx.doi.org/10.1080/23335432.2019.1621205, PMID: 34042002.

[14] M. Fonseca, M. Bergere, J. Candido, F. Leboeuf, R. Dumas, S. Armand, The conventional gait model's sensitivity to lower-limb marker placement, Sci. Rep. 12 (1) (2022) 1–8.

[15] I. Gaboury, M. Tousignant, H. Corriveau, M. Menear, G. Le Dorze, C. Rochefort, B. Vachon, A. Rochette, S. Gosselin, F. Michaud, J. Bollen, S. Dean, Effects of Telerehabilitation on patient adherence to a rehabilitation plan: Protocol for a mixed methods trial, JMIR Res. Protocols (ISSN: 19290748) 10 (10) (2021) http://dx.doi.org/10.2196/32134.

[16] A. Nuara, M. Fabbri-Destro, E. Scalona, S.E. Lenzi, G. Rizzolatti, P. Avanzini, Telerehabilitation in response to constrained physical distance: an opportunity to rethink neurorehabilitative routines, J. Neurol. 269 (2) (2022) 627–638.

[17] E. Martini, M. Boldo, S. Aldegheri, N. Valè, M. Filippetti, N. Smania, M. Bertucco, A. Picelli, N. Bombieri, Enabling Gait Analysis in the Telemedicine Practice through Portable and Accurate 3D Human Pose Estimation, Comput. Methods Programs Biomed. (ISSN: 01692607) 225 (2022) 107016, http://dx.doi.org/10.1016/j.cmpb.2022.107016, URL https://linkinghub.elsevier.com/retrieve/pii/S0169260722003984.

[18] C.E. Levy, E. Silverman, H. Jia, M. Geiss, D. Omura, Effects of physical therapy delivery via home video telerehabilitation on functional and health-related quality of life outcomes, J. Rehabil. Res. Dev. (ISSN: 19381352) 52 (3) (2015) 361–370, http://dx.doi.org/10.1682/JRRD.2014.10.0239.

[19] Z. Cao, T. Simon, S.E. Wei, Y. Sheikh, Realtime multi-person 2d pose estimation using part affinity fields, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7291–7299.

[20] A. Ong, I.S. Harris, J. Hamill, The efficacy of a video-based marker-less tracking system for gait analysis, Comput. Methods Biomech. Biomed. Eng. 20 (10) (2017) 1089–1095.

[21] R. Tanaka, H. Takimoto, T. Yamasaki, A. Higashi, Validity of time series kinematical data as measured by a markerless motion capture system on a flatland for gait assessment, J. Biomech. (ISSN: 18732380) 71 (2018) 281–285, http://dx.doi.org/10.1016/j.jbiomech.2018.01.035.

[22] E. D'Antonio, J. Taborri, I. Mileti, S. Rossi, F. Patane, Validation of a 3D Markerless System for Gait Analysis Based on OpenPose and Two RGB Webcams, IEEE Sens. J. (ISSN: 15581748) 21 (15) (2021) 17064–17075, http://dx.doi.org/10.1109/JSEN.2021.3081188.

[23] M. Tousignant, H. Moffet, S. Nadeau, C. Mérette, P. Boissy, H. Corriveau, F. Marquis, F. Cabana, P. Ranger, É.L. Belzile, R. Dimentberg, Cost analysis of in-home telerehabilitation for post-knee arthroplasty, J. Med. Internet Res. (ISSN: 14388871) 17 (3) (2015) http://dx.doi.org/10.2196/jmir.3844.

[24] R. Hwang, N.R. Morris, A. Mandrusiak, J. Bruning, R. Peters, D. Korczyk, T. Russell, Cost-Utility Analysis of Home-based Telerehabilitation Compared with Centre-based Rehabilitation in Patients with Heart Failure, Heart Lung Circ. (ISSN: 14442892) (2018) http://dx.doi.org/10.1016/j.hlc.2018.11.010.

[25] G. Maresca, M.G. Maggio, R. De Luca, A. Manuli, P. Tonin, L. Pignolo, R.S. Calabrò, Tele-Neuro-Rehabilitation in Italy: State of the Art and Future Perspectives, Front. Neurol. (ISSN: 16642295) 11 (2020) http://dx.doi.org/10.3389/fneur.2020.563375.

[26] M.G. Maggio, M. Foti Cuzzola, P. Calatozzo, D. Marchese, A. Andaloro, R.S. Calabrò, Improving cognitive functions in adolescents with learning difficulties: A feasibility study on the potential use of telerehabilitation during Covid-19 pandemic in Italy, J. Adolesc. (ISSN: 10959254) 89 (2021) 194–202, http://dx.doi.org/10.1016/j.adolescence.2021.05.005.

[27] G. Milani, G. Demattè, M. Ferioli, G. Dallagà, S. Lavezzi, N. Basaglia, S. Straudi, Telerehabiltation in Italy During the COVID-19 Lockdown: A Feasibility and Acceptability Study, Int. J. Telerehabil. 13 (1) (2021) http://dx.doi.org/10.5195/ijt.2021.6334.

[28] I.E. Mosca, E. Salvadori, F. Gerli, L. Fabbri, S. Pancani, G. Lucidi, G. Lombardi, L. Bocchi, S. Pazzi, F. Baglio, F. Vannetti, S. Sorbi, C. Macchi, Analysis of Feasibility, Adherence, and Appreciation of a Newly Developed Tele-Rehabilitation Program for People With MCI and VCI, Front. Neurol. (ISSN: 16642295) 11 (2020) http://dx.doi.org/10.3389/fneur.2020.583368.

[29] M.G. Benedetti, E. Beghi, A. De Tanti, A. Cappozzo, N. Basaglia, A.G. Cutti, A. Cereatti, R. Stagni, F. Verdini, M. Manca, et al., SIAMOC position paper on gait analysis in clinical practice: General requirements, methods and appropriateness. Results of an Italian consensus conference, Gait Posture 58 (2017) 252–260.

[30] R. Di Marco, S. Rossi, E. Castelli, F. Patanè, C. Mazzà, P. Cappa, Effects of the calibration procedure on the metrological performances of stereophotogrammetric systems for human movement analysis, Measurement 101 (2017) 265–271.

[31] R.M. Kanko, E.K. Laende, G. Strutzenberger, M. Brown, W.S. Selbie, V. DePaul, S.H. Scott, K.J. Deluzio, Assessment of spatiotemporal gait parameters using a deep learning algorithm-based markerless motion capture system, J. Biomech. 122 (2021) 110414, http://dx.doi.org/10.1016/j.jbiomech.2021.110414.

[32] R.B. Davis, III, S. Ounpuu, D. Tyburski, J.R. Gage, A gait analysis data collection and reduction technique, Hum. Movement Sci. 10 (5) (1991) 575–587.

[33] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, Y.A. Sheikh, OpenPose: Realtime multi-person 2D pose estimation using part affinity fields, IEEE Trans. Pattern Anal. Mach. Intell. (2019).

[34] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, Microsoft coco: Common objects in context, in: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, Springer, 2014, pp. 740–755.

[35] T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, Microsoft coco: Common objects in context – version 3, https://arxiv.org/abs/1405.0312, Updated: 2015-02-21, http://dx.doi.org/10.48550/arXiv.1405.0312.

[36] M. Andriluka, L. Pishchulin, P. Gehler, B. Schiele, 2D human pose estimation: New benchmark and state of the art analysis, in: Procedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, pp. 3686–3693, http://dx.doi.org/10.1109/CVPR.2014.471.

[37] H.J. Woltring, A fortran package for generalized, cross-validatory spline smoothing and differentiation, Adv. Eng. Softw. (1978) 8 (2) (1986) 104–113.

[38] J. Perry, Gait Analysis. Normal and Pathological Function, SLACK Incorporated, Downey, CA, ISBN: 978-1-55642-192-1, 1992, p. 524.

[39] A. Middleton, S.L. Fritz, M. Lusardi, Walking speed: The functional vital sign, J. Aging Phys. Activity 23 (2) (2015) 314–322, http://dx.doi.org/10.1123/japa.2013-0236.

[40] T.C. Pataky, One-dimensional statistical parametric mapping in python, Comput. Methods Biomech. Biomed. Eng. 15 (3) (2012) 295–301.

[41] P. Mishra, C.M. Pandey, U. Singh, A. Gupta, C. Sahu, A. Keshri, Descriptive statistics and normality tests for statistical data, Ann. Card. Anaesth. 22 (1) (2019) 67.

[42] RStudio Team, RStudio: Integrated Development Environment for R, RStudio, PBC, Boston, MA, 2021, URL https://www.rstudio.com.

[43] A. Ferrari, M.G. Benedetti, E. Pavan, C. Frigo, D. Bettinelli, M. Rabuffetti, P. Crenna, A. Leardini, Quantitative comparison of five current protocols in gait analysis, Gait Posture (ISSN: 0966-6362) 28 (2) (2008) 207–216, http://dx.doi.org/10.1016/j.gaitpost.2007.11.009, URL http://www.ncbi.nlm.nih.gov/pubmed/18206374,

[44] R. Di Marco, S. Rossi, V. Racic, P. Cappa, C. Mazzà, Concurrent repeatability and reproducibility analyses of four marker placement protocols for the foot-ankle complex, J. Biomech. (ISSN: 00219290) 49 (14) (2016) 3168–3176, http://dx.doi.org/10.1016/j.jbiomech.2016.07.041, URL http://dx.doi.org/10.1016/j.jbiomech.2016.07.041.

[45] R. Di Marco, E. Scalona, A. Pacilli, P. Cappa, C. Mazzà, S. Rossi, How to choose and interpret similarity indices to quantify the variability in gait joint kinematics, Int. Biomech. 5 (1) (2018) 1–8.

[46] JCGM, International vocabulary of metrology – Basic and general concepts and associated terms (VIM) 3rd edition Vocabulaire international de métrologie – Concepts fondamentaux et généraux et termes associés (VIM) 3 e édition, (Vim) 2012.

[47] M. Iosa, A. Cereatti, A. Merlo, I. Campanini, S. Paolucci, A. Cappozzo, Assessment of waveform similarity in clinical gait data: the linear fit method, BioMed Res. Int. 2014 (2014).

[48] J.M. Bland, D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 327 (8476) (1986) 307–310.

[49] J.M. Bland, D.G. Altman, Measuring agreement in method comparison studies, Stat. Methods Med. Res. 8 (2) (1999) 135–160.

[50] A.S. Albahri, J.K. Alwan, Z.K. Taha, S.F. Ismail, R.A. Hamid, A. Zaidan, O.S. Albahri, B. Zaidan, A.H. Alamoodi, M. Alsalem, IoT-based telemedicine for disease prevention and health promotion: State-of-the-art, J. Netw. Comput. Appl. 173 (2021) 102873.

[51] M. Mijwil, R. Doshi, K.K. Hiran, A.-H. Al-Mistarehi, M. Gök, Cybersecurity challenges in smart cities: An overview and future prospects, Mesop. J. Cybersecur. 2022 (2022) 1–4.

[52] L. Keselman, J. Iselin Woodfill, A. Grunnet-Jepsen, A. Bhowmik, Intel realsense stereoscopic depth cameras, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 1–10.

[53] I. de Medeiros Esper, O. Smolkin, M. Manko, A. Popov, P.J. From, A. Mason, Evaluation of RGB-D multi-camera pose estimation for 3D reconstruction, Appl. Sci. 12 (9) (2022) 4134.