Evaluation of the intra- and inter-observer reproducibility of a method for studying three-dimensional knee kinematics

David R. Labbe, Nicola Hagemeister, Mathieu Charbonneau, Jacques A. de Guise
Laboratoire de recherche en imagerie et orthopédie (LIO)
Centre hospitalier de l'Université de Montréal (CHUM)
Montréal, Canada
lio@etsmtl.ca

***Abstract*— A major difficulty in evaluating in-vivo knee joint kinematics during walking is the artifacts created by movement of the skin and soft tissue relative to the underlying bones. An attachment system called exoskeleton has been proposed to remedy this problem; its precision is known as well as the repeatability of its calibration method. The objective of this study was to verify the intra- and inter-observer repeatability of this system. Also, as the installation is more complex than that of other systems, the effect of the observers' experience was investigated. Results show a mean repeatability below 1.3° in the intra- and inter-observer settings. Observer experience had no effect on the results although a slight learning curve may exist for novice observers.**

Keywords-component; three-dimensional kinematics, knee kinematics, anatomical frames, functional method, validation, repeatability

I. INTRODUCTION

In-vivo analysis of knee joint kinematics during walking is a difficult task; the difficulty mostly arises from movement of the skin and soft tissue relative to underlying bones. When using surface mounted markers, the error incurred can be anywhere from 2 to 20 times that of the equipment used to track these markers [1]. Different attachment systems have been proposed in an attempt to limit this artifact and follow bone movement with greater precision.

One such attachment system is an exoskeleton proposed by Sati et al. [2] which is composed of a tibial component as well as a harness that fixes to the femur. The precision of this method in following underlying bone movement has been documented by fluoroscopic study [3]; the calibration used with this system has also been shown to be reproducible [4]. What is not known however is the repeatability of the measure between installations of the femoral and tibial components as well as the variability introduced when the evaluation is performed by different observers.

The objective of this study is therefore to evaluate the intra- and inter-observer repeatability of the aforementioned exoskeleton and to establish if there exists a learning curve (i.e. are measurements more reproducible when taken by an experienced observer).

II. MATERIALS AND METHODS

A. Subjects

Fifteen subjects (9 men and 6 women) participated in this study; their mean age was 27.13 years (23 to 36 years old). None of the subjects had a history of any type of musculoskeletal problem and they all gave their written consent by signing the appropriate forms which were approved by the institutional ethics committees. Subjects walked on a treadmill (ADAL-COP, TECMachine, France) at a comfortable speed while kinematics were recorded using a 120 Hz 3D digital optical system (Vicon, Oxford Metrics Ltd., Oxford, England). The system utilized 6 infrared cameras positioned around the treadmill with 4 of the cameras recording the trajectories of markers placed on the right side of the subject.

The treadmill was equipped with two force plate-forms allowing for the measurement of 3D ground reaction forces which were used to divide the recordings into distinct gait cycles.

B. Experimental protocol

Three observers participated in this study: 1 expert observer who had installed the harness on over 300 knees, 1 initiated observer who had installed it on approximately 20 knees and 1 novice observer who had only minimal training for installing the harness and had practiced on 3 different knees. During the course of the experiment, each of these observers performed the installation of the harness twice on each subject. The order of the observers' installations was chosen at random before each session.

When the subject first arrived, he was fitted with a pair of sandals used for every subject so as to eliminate any variations originating from different footwear. He was then asked to walk on the treadmill in a regular manner for a warm-up period

---

Funding provided by NSERC, CIHR and Canada Research Chair in 3D Imaging and Biomedical Engineering

of 10 minutes during which his comfortable walking speed was established. The treadmill was initially set to a speed calculated based on the subject's leg length and was then adjusted by visually assessing the subject's apparent effort or upon his request; established comfortable walking speed ranged from 0.92 to 1.22 m/s for all subjects.

Reflective markers were affixed to the malleoli of the subject's right lower limb. A belt with reflective markers was tightly apposed around his waist, directly over the iliac spines. The subject was then equipped with the harness as well as with the tibial component, both of which were mounted with markers (Fig. 1). He resumed walking at his determined comfortable speed; after a period of 3 minutes, three 20 second recordings of his kinetics and kinematics were taken. Following these recordings, the function postural (FP) calibration method [4] was performed in order to determine joint centers and to align the axes. This entire procedure was repeated twice by each observer.

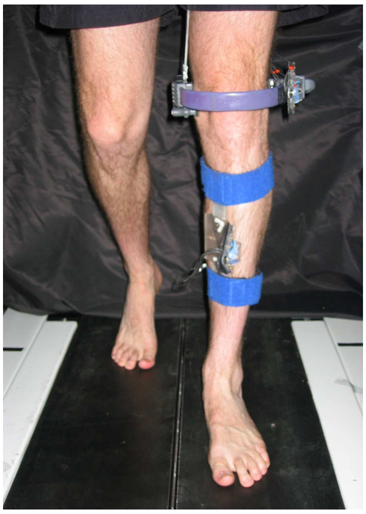

Figure 1. Femoral and tibial components of the exoskeleton

C. Data analysis

The FP method calibration data were combined with the gait data recorded. Using the method described by Grood and Suntay [5], 6 sets of kinematic parameters (3 observers x 2 installations) were computed for each subject. Each set was composed of flexion/extension, abduction/adduction and internal/external tibial rotation.

Kinematic data were divided into individual gait cycles using foot contact patterns as obtained from the right-sided force plate-form. All cycles were normalized to 100 sample points representing 100% of the gait cycle. Using a visual interface, all flexion, abduction, tibial rotation and vertical force curves for a given installation were presented with a representation of the mean curve and 2 standard deviations. Curves exceeding these 2 standard deviations where excluded from data analysis as they were judged to be a result of bad cycle division.

Repeatability

To evaluate the repeatability of the measures, all gait cycles from a single installation were grouped together. Also, in order to evaluate the evolution of repeatability during the study, all gait cycles for a single subject were grouped together regardless of observer. For each kinematic parameter, the mean deviation from the average curve was calculated for the inter- and intra-observer settings.

Reliability

For further data analysis, 15 points of interest were considered [6, 7]. The first 5 points, individually selected from the vertical force curve of every gait cycle, represent the point of contact of the right foot with the force plate-form (Fig. 2a, PT1), the point were the foot leaves the plate-form (Fig. 2a, PT5) as well as two local maximums (Fig . 2a, PT2, PT4) and one local minimum (Fig. 2a, PT3). The next four points are selected from the antero-posterior force curve and represent two local maximums (Fig. 2b, PT6, PT9), one local minimum (Fig. 2b, PT7) and the root value of the curve between points 7 and 9 (Fig. 2b, PT8). Using the medio-lateral force curve, two local minimums (Fig. 2c, PT10, PT13), two local maximums (Fig. 2c, PT12, PT14) and the root of the curve between points 10 and 12 (Fig. 2c, PT11) were selected. Finally, the point of maximal knee flexion was also considered.

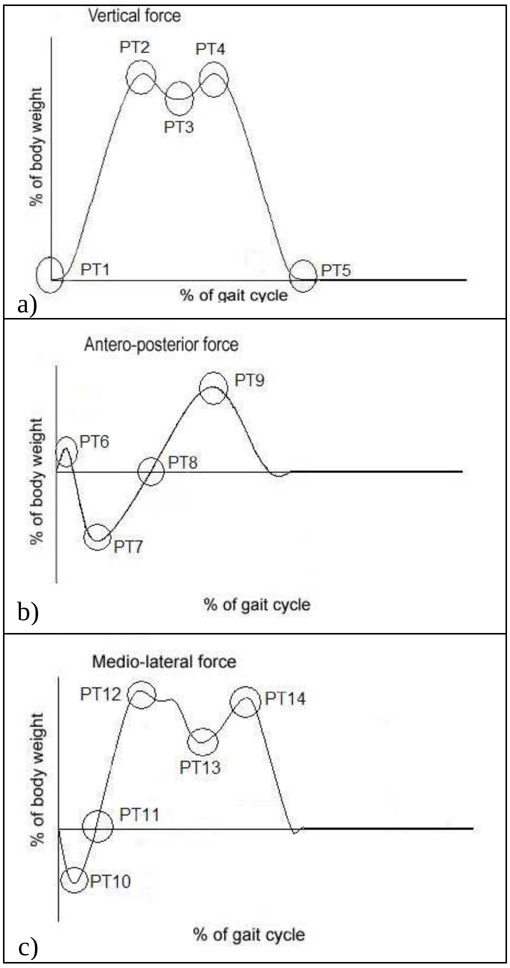

Figure 2. Selected points of interest (adapted from [6, 7])

For each kinematic parameter, intra-class correlation (ICC) was calculated for each of the 15 points of interest in the intra- and inter-observer settings.

III. RESULTS

The data from 12 of the initial 15 subjects were analyzed. Incomplete optical recordings for subjects 12 through 14 rendered their data unusable for the current study.

Repeatability

In the intra-observer setting, repeatability values for all subjects ranged from 0.8° to 1.3° for flexion/extension. For abduction/adduction and tibial rotation, the values ranged from 0.8 to 1.0° and from 0.9 to 1.1° respectively (Table 1).

TABLE I. INTRA-OBSERVER REPEATABILITY VALUES FOR ALL OBSERVERS

Observer	**Flex/ext (°)**	**Abd/add (°)**	**Int/ext rot. (°)**
#1	1.3±0.8	0.8±0.8	1.0±0.5
#2	0.8±0.6	1.0±0.6	0.9±0.5
#3	0.9±0.5	1.0±0.5	1.1±1.3

Flex/ext = flexion/extension, Abd/add = abduction/adduction, Int/ext rot. = internal/external tibial rotation.

Inter-observer repeatability values were similar with mean values of 1.3° for flexion/extension, 0.7° for abduction/adduction and 1.0 ° for tibial rotation (Table II).

TABLE II. MEAN INTER-OBSERVER REPEATABILITY VALUES

Flex/ext (°)	**Abd/add (°)**	**Int/ext rot (°)**
1.3±0.4	0.7±0.2	1.0±0.6

Abbreviations as in Table I.

Fig. 3 shows the intra-observer repeatability values obtained for each observer and for each kinematic parameter. Overall, average intra-observer deviation for flexion/extension across all observers for the first 6 subjects was 1.19° compared with 0.90° for the last 6 subjects. The expert observer showed a slight improvement in intra-observer repeatability throughout the study while the initiated observer showed a slight decrease in repeatability. The largest improvement in intra-observer repeatability was showed by the novice observer (Fig. 4) although this improvement was not significant.

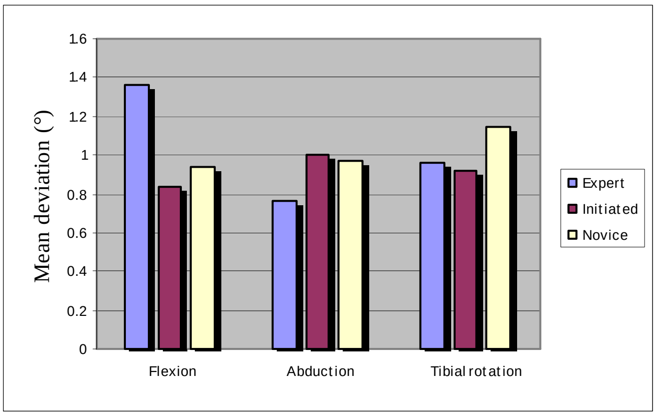

Figure 3. Intra-observer mean repeatability values for all kinematic parameters

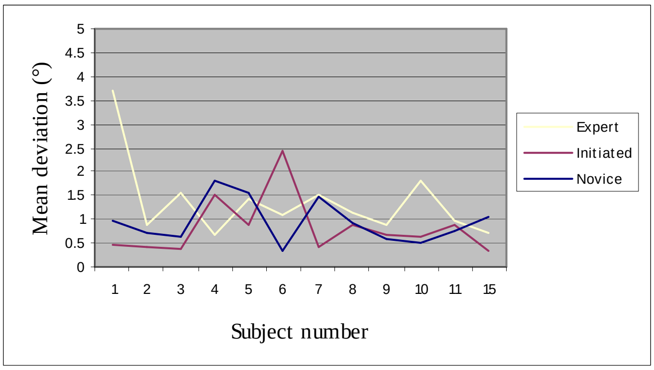

Figure 4. Flexion/extension mean deviations for all observers

Reliability

Mean ICC values across all 15 points of interest are shown in table III for inter- and intra-observer settings. For knee flexion, inter observer ICC values ranged from 0.90 to 0.97 and from 0.89 to 0.97 for intra-observer. In knee abduction and tibial rotation, those values ranged from 0.77 to 0.98 for both settings.

TABLE III. MEAN ICC VALUES FOR 15 POINTS OF INTEREST

	Flex/ext(°)	**Abd/add(°)**	**Int/ext rot(°)**
Inter	0.94±0.02	0.91±0.05	0.88±0.05
Intra	0.94±0.02	0.91±0.05	0.88±0.05

IV. DISCUSSION

In this study, we evaluated ICC and mean deviation values in an intra- and inter-observer context. ICC values give us information concerning the correlation of the different curves whereas mean deviations are more telling of the repeatability, in other words the actual range of values. The high ICC values obtained indicate high correlation of the curves while the mean deviations are lower than 10% of the range of values for all parameters and settings.

Our results show mean deviation values that are a little higher than those reported by Hagemeister et al. [4]. Indeed, the values reported in that study for the intra-observer setting were slightly lower then our findings: below 0.8° for flexion/extension, 1.2° for abduction/adduction and 0.9° for tibial rotation. However, the maximum inter-observer value for all kinematic parameters was 0.4° whereas we found a value of 1.3° for flexion/extension.

This difference is not surprising since Hagemeister et al. evaluated only the repeatability of the FP calibration method which is used with the exoskeleton. In other words, no reinstallation of the harness or tibial component was done between recordings. It is therefore probable that the higher deviations reported here, particularly in the inter-observer setting, are due to reinstallation. The small difference between deviations reported in both studies indicates that while reinstallation of the harness does introduce an added variability, it is minimal and less important than that accounted for by the calibration method itself.

Similar values for the mean deviations in the intra- and inter-observer contexts indicate that reinstallation of the exoskeleton by different observers does not introduce more variation than repeated installations by a single observer. Furthermore, the ICC values obtained in both settings show no difference between intra- and inter-observer repeatability.

Using intra-observer mean deviations, no clear distinction can be made between observers of different experience levels. In fact, it can be observed in Fig. 3 that for flexion/extension the expert observer produced the highest mean deviation. This suggests that even a novice observer can obtain repeatable results with little training. It should be noted however that although little experience is needed for installation of the exoskeleton, adequate training is required to assure that the femoral harness is affixed to the limb in the proper manner.

Despite the absence of a difference in mean values for observers with different experience levels, a slight improvement in intra-observer repeatability can be seen through time for the novice observer (Fig. 4). This improvement is not apparent for the initiated observer and is negligible for the expert observer. This learning curve is relatively small but indicates nonetheless that some improvement may be seen over time in novice observers. A greater number of observers would be necessary to confirm this hypothesis.

The results obtained indicate that reinstallation of the exoskeleton introduces a greater variability than simply repeating the calibration phase; this is to be expected. Furthermore, repeatability is equal regardless of whether multiple installations are performed by a single observer or not. The results also show that the observer's experience has little to no effect on repeatability. These conclusions are somewhat limited by the fact that the data of only 12 subjects was used and even more so by the fact that a single observer of each experience level participated. Also, intra-observer results are limited as only two installations were performed by each observer on each subject.

Day to day repeatability is yet to be tested for the exoskeleton and should be to allow for comparison of data recorded on different days.

Once fully validated, the system could be used to assess the effect of treatment of different pathologies such as osteoarthritis, ligament ruptures, etc. To that effect, a portable version of this 3D kinematics recording system has been developed for use in a clinical setting using an electromagnetic device instead of the Vicon digital optical system.

ACKNOWLEDGMENT

The authors thank Karine Pelletier, Gerald Parent and Sabine Husse for their contribution to this study.

REFERENCES

[1] M. Sati, J.A. de Guise, S. Larouche and G. Drouin, “Quantitative assessment of skin-bone movement at the knee”. The Knee, 1996. **3**(3): p. 121-138.

[2] M. Sati, J.A. de Guise, S. Larouche and G. Drouin, “Improving in vivo knee kinematic measurements: application to prosthetic ligament analysis”. The Knee, 1996. **3**(4): p. 179-190

[3] S. Ganjikia, N. Duval, L. Yahia and J. de Guise, “Three-dimensional knee analyzer validation by simple fluoroscopic study”. The Knee, 2000. **7**(4): p. 221-231.

[4] N. Hagemeister, G. Parent, M. Van de Putte, N. St-Onge, N. Duval and J. de Guise, “A reproducible method for studying three-dimensional knee kinematics”. J Biomech, 2005. **38**(9): p. 1926-31.

[5] E.S. Grood and W.J. Suntay, “A joint coordinate system for the clinical description of three-dimensional motions: application to the knee”. J Biomech Eng, 1983. **105**(2): p. 136-44.

[6] E.Y. Chao, R.K. Laughman, E. Schneider and R.N. Stauffer, “Normative data of knee joint motion and ground reaction forces in adult level walking”. J Biomech, 1983. **16**(3): p. 219-33.

[7] M.G. Benedetti, F. Catani, A. Leardini, E. Pignotti and S. Giannini, “Data management in gait analysis for clinical applications”. Clin Biomech (Bristol, Avon), 1998. **13**(3): p. 204-215.