Case study of surgeon's kinematics performing arthroscopy in real and simulation scenarios: a multisensor approach

Nicola Abeni
DIMI, University of Brescia
Brescia, Italy
nicola.abeni@unibs.it

Emilia Scalona
DSMC, University of Brescia
Brescia, Italy
emilia.scalona@unibs.it

Marco Ghidelli
DSMC, University of Brescia
Brescia, Italy
marco.ghidelli@unibs.it

Maristella F. Saccomanno
DSMC, dept. of Bone and Joint Surgery
University of Brescia, Spedali Civili
Brescia, Italy
maristella.saccomanno@unibs.it

Giuseppe Milano
DSMC, dept. of Bone and Joint Surgery
University of Brescia, Spedali Civili
Brescia, Italy
giuseppe.milano@unibs.it

Matteo Lancini
DSMC, University of Brescia
Brescia, Italy
matteo.lancini@unibs.it

***Abstract*—Surgical competence involves a spectrum of skills crucial for surgical trainees, however, the formal study of operative teaching and assessment lags other aspects of surgical education. Arthroscopy presents technical challenges requiring precise coordination and spatial awareness. The present study aims to study the feasibility of bridging the gap between the operating room and simulation environments. This has been achieved by comparing kinematic measurements during anterior cruciate ligament (ACL) reconstruction surgeries performed by a surgeon in both settings. Using inertial sensors and depth cameras, kinematic metrics were recorded and analyzed for eight surgical stages replicable across the two mentioned environments. The study established a strict protocol for operating room access and assessed the usability of sensor equipment. This work showed that it is possible to obtain kinematics metrics through IMUs inside the operating room, even if some discrepancies are present with the training environment on a knee mannequin. It has also been investigated the possibility of using an Azure Kinect DK as an alternative to the IMUs, but some criticalities have been highlighted.**

Keywords—Surgical competence, arthroscopy, kinematic measurement, IMUs, skeletonization

I. Introduction

Surgical competence encompasses knowledge, technical skills, decision-making abilities, communication skills, and leadership capabilities. Technical proficiency is crucial for surgical trainees, as it serves as quality assurance for their future practice [1]. However, the formal study of operative teaching and testing has experienced slower development compared to other aspects of surgical education [2]. Moreover, in recent years, factors such as the European Working Time Directive and financial pressures have reduced the opportunity for trainees to learn surgical skills, posing a risk that some surgeons may not achieve sufficient skill levels upon completing their training [3]. The increasing attention of the public and media towards the performance of doctors has sparked an interest in the development of robust methods for assessing technical skills [4].

Among the diverse surgery techniques, arthroscopy is technically challenging because it requires good visual-spatial coordination to control tools and translate three-dimensional structures into two-dimensional images. This is typically learned through a step-by-step apprenticeship approach in the operating scenario, which is time and money-inefficient and linked to patient iatrogenic damage [5]. To address deficiencies, the emphasis is shifting to outside-the-operating-room skill learning and assessment tools [6]. However, this assessment is largely subjective and lacks objectivity. Therefore, to introduce quantitative parameters, the idea of this preliminary study is to investigate the existing differences between the operating room and the simulation scenario performing kinematics measurements on a surgeon in both environments. The surgical procedure considered is the anterior cruciate ligament (ACL) reconstruction; regarding the simulation scenario, the surgery is reproduced on a knee mannequin.

Two different approaches have been adopted: the first one consists of the use of inertial sensors (Xsens MTw Awinda, The Netherlands), while the other is based on the human body skeletonization through RGB depth cameras (Microsoft Azure Kinect DK).

A. Related works

In the last few years, diverse studies have explored assessing resident proficiency in arthroscopic procedures, primarily through simulation methods like virtual reality and cadaveric models. Tools like the Arthroscopic Surgical Skill Evaluation Tool (ASSET) [7] exist; being checklists, they rely on evaluation based on experience, but they remain still subjective. The interest is so shifting towards the collection of quantitative data, for example using motion analysis systems. Some of them, as presented in [8] and [9] are suitable just for the simulation scenario because they rely on wired connections and electromagnetic tracking technology, which would impede the correct proceedings of the operation. Other studies [10] and [11] propose instead the use of wearables sensors, but they are just a few, suggesting a deeper investigation.

The various studies shared a common focus on kinematic metrics, particularly those concerning shoulder and elbow motion, which were analyzed to determine the angular range of motion across the three biomechanical planes. They have considered the shoulder add/abduction and flex/extension, the elbow flex/extension and prono/supination.

Another noteworthy parameter considered was the total path length, representing the trajectory covered by surgical instruments. While virtual simulators, such as those discussed in [10], can directly measure this parameter, the present study approximated it using the hand trajectory length due to the simulation being conducted with a knee mannequin.

In both [10] and [11] the focus was on understanding how the mentioned metrics can be useful to differentiate between different dexterity levels among expert, intermediate and novice surgeons. They found that experience could be identified through kinematic metrics. However, neither study explored the similarities between real operating rooms and simulation scenarios.

This is where the present study comes in. The novelty of the present work lies mainly in the following aspect: the conduction of measurements directly inside the operating room in a real surgery. Working in a scenario like that it's critical and implies a strict access protocol and the use of non-invasive instruments for the surgeon while operating.

II. Methodology

The measurement setup consisted of nine inertial measurement units (Xsens MTw Awinda, The Netherlands) and a depth camera (Microsoft Azure Kinect DK) running a skeletonization algorithm. The selection of IMUs was based on their non-invasive nature, as determined through feedback from surgeons via a modified version of the System Usability Scale. The Azure Kinect DK camera was added to explore the potential of gathering data without requiring the surgeon to wear sensors.

A. IMU sensors

The Xsens MTw Awinda are wireless sensors with a communication range of up to 20 meters, a battery life of 6 hours and can operate in a temperature range of 0°C to 50°C. Each sensor weighs 16 grams and has a dimension of 47x30x13 millimeters. Furthermore, they sample data at 100 Hz which is a frequency suited for human movements.

As our kinematic analysis targets the shoulders and elbows, we opted to monitor the upper body of the surgeon. Following recommendations from the manufacturer, we adopted the biomechanics model called "Upper Body no Hands". To achieve this, the sensors must be positioned on the surgeon's body with the configuration illustrated in Figure 1.

Fig. 1. IMU sensors positioning on the upper body of the surgeon. In yellow are highlighted the sensors that should be positioned near the wrist, but since the surgeon wore sterile gloves were positioned near the elbow.

After performing the calibration, we gathered kinematic data with the motion capture software MVN Analyze, 2023, version 2023.2.

B. Depth Camera and Skeletonization

The depth camera used in this study was the Kinect Azure DK, offering synchronized RGB and depth video. We selected an RGB video resolution of 1920x1080 pixels and a depth video resolution of 640x576 pixels. The acquisition was conducted at 30 frames per second, with exposure manually set to predefined values to accommodate lighting conditions. The camera was positioned in front of the surgeon as carefully as possible to avoid obstructing the surgical staff and to prevent interference with sterile materials.

Microsoft provides an SDK for tracking multiple bodies in the 3D space; each body consists of a series of keypoints connected with segments to reconstruct a skeletonized version of the human body. The body tracking SDK associates an ID to each body identified in the scene. This is not always done in the right manner since sometimes the same ID is associated with a different body and so we lose the reference to the surgeon which is the only skeleton in which we have an interest. To solve the problem of misrecognition between bodies' ID, we developed an algorithm that remaps the bodies relying on the position of the neck keypoint: frame by frame we analyze this position and associate the same ID to the skeleton which has the nearest neck keypoint to the previous frame. Then we plot the different skeletons and manually select the surgeon by comparing the graphs to the video.

Once the surgeon is identified and the 3D position of the keypoints has been extracted, we used the software OpenSim [12] to perform the inverse kinematics and compute the joint angles. The OpenSim model used to accomplish this task is the Rajagopal 2015 [13] appropriately modified, and the keypoints used from the skeletonization are the following: pelvis, spine chest, left and right shoulders, left and right elbows and left and right wrist. The inverse kinematic tool provided by OpenSim analyzes each frame of motion, calculating generalized coordinates value to position the model in a pose that closely aligns with experimental keypoints and coordinate data. This alignment is achieved by solving a weighted least squares problem, aiming to minimize errors in both keypoints and coordinates.

Fig. 2. Azure body skeletonization SDK inference on a depth frame from the Azure Kinect DK. The scenario is the operating room and it's possible to notice the position of the image segmentation that identifies different people and the skeleton key points on each body.

C. Experiment protocol

In this preliminary feasibility assessment phase, the primary surgeon from the Department of Bone and Joint Surgery at Brescia Hospital was selected as the subject of the experiment. This choice ensures that we gather data from an expert source, which can then serve as a reference for surgeons with varying levels of expertise. The same subject performed the surgical procedure once in the operating room and once in the simulation environment.

A strict access protocol has been defined for measurements conducted in the operating room, ensuring adherence to all safety and sterilization recommendations. First, the surgeon's anthropometric measurements are taken, which are crucial for calibrating the IMU sensors. It's important to note that this procedure is only necessary the first time or when recordings involve a new surgeon. Then, the surgeon is equipped with the IMU sensors, and calibration takes place. Subsequently, the surgeon enters the operating room to perform the surgery. While two individuals are permitted to enter the operating room to capture backup videos and record the surgery using the Azure Kinect DK, one person remains outside the operating room to monitor the connection between the IMUs and the laptop, which serves as a receiver.

In the simulation, we aimed to replicate surgical procedures as closely as possible, without requiring adherence to a strict protocol. There is no need to wear sterile clothing, to pay particular attention to other people assisting the surgeon and no hitches can happen. So, we set up a room with a knee mannequin and all the necessary surgical instruments. Then the surgeon is equipped with the IMU sensors, and the depth camera starts the recording.

A real surgical procedure includes a lot of parts or movements which are not standard because they depend on the patient's ligament level of damage, on the operated knee (left or right), and on the experience of the staff assisting the surgeon and cannot be replied to in simulation. The simulation tasks found in other research do not replicate the real tasks but just try to replicate some common movements. During the co-design phase with surgeons, we conducted a brief analysis to determine which stages should be recorded in the operating room and subsequently replicated in the simulation. Table 1 presents the selected surgical stages.

TABLE I. SURGICAL STAGES REPLICABLE IN SIMULATION

Surgical stages	Surgical gestures
Femoral tunnel	1. Femoral pointer positioning (FP)
	2. Femoral tunnel creation (FC)
	3. Traction thread insertion (TI)
Tibial tunnel	4. Tibial pointer positioning (TP)
	5. Tibial tunnel creation (TC)
	6. Traction thread taking (TT)
Graft insertion	7. Graft positioning (GP)
	8. Graft fixation (GF)

D. Data organization and synchronization

The two datasets have been synchronized using the timestamp and then segmented into eight stages according to those reported in Table 1. Finally, the range of motion has been computed for both Azure Kinect and Xsens data in each stage and allows the comparisons between the two instruments and between the operating room and the simulation scenario. The range of motion is determined by calculating the difference between the 95th and 5th percentiles for the joint rotational position for each gesture.

III. RESULTS

A. System usability scale

To understand whether the sensor equipment created discomfort or obstacles, a questionnaire was administered to the surgeon. The results of this questionnaire are summed up in Table 2.

TABLE II. SYSTEM USABILITY SCALE

	Questions		
	Do you think it might cause discomfort during the surgery?	***Do you think it may affect your freedom of movement?***	***Do you think there is a risk that the device may fall?***
Upper t-shirt for shoulder and stern sensors	NO	NO	NO
Band for upper arm sensors	NO	NO	NO
Band for wrist sensors	YES[a]	NO	NO

[a] The wrist sensors are not very uncomfortable, but the surgeon must wear gloves, the solution adopted is to shift the sensors on the forearms since the Xsens software has performed a good calibration despite this change of position.

The surgeon spontaneously reported that the whole sensor setup can cause excessive heat and sweat, especially in the warmer months of the year, but this does not impede the use of this setup and does not alter the correct unfolding of the surgical procedure.

B. Kinematic results

The following results refer to a single measurement on the same surgeon performing knee arthroscopy in both the operating room and on a knee mannequin. Since a single test does not bring statistical relevance these are just preliminary results to understand the feasibility of the proposed measurement methodology and to qualitatively understand differences between the simulation and the real surgery.

Figures 3 to 8 detail the range of motion (ROM) concerning elbows and shoulders for each of the previously mentioned eight stages. Each figure will display two graphs, presenting both Azure and Xsens data in both real and simulation scenarios. In each graph, the blue curve represents the Xsens data, while the orange one represents the Azure data. These results do not show the intra/extra rotation for the shoulder and the prono/supination for the elbow because the skeletonizer approximates the body segment as one-dimensional rigid rods, so we do not have any information about that degree of freedom. Those angles could just be estimated but this approximation is not reliable for the comparison with the IMUs data.

1) Right shoulder range of motion

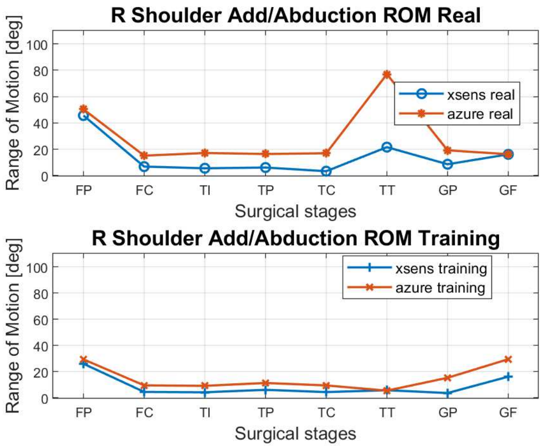

Fig. 3. Right shoulder add/abduction range of motion in both operating room (above) and simulation scenario (below). Data from Azure Kinect are represented in orange while data from Xsens in blue.

The differences between ROMs coming from Xsens and Azure in the right shoulder add/abduction in the operating room range between a minimum of 0 degrees in the GF phase to a maximum of 13 degrees in the TC phase, without considering the peak at 80 degrees for the TT phase. Regarding the simulation environment, the differences in terms of ROMs between the two instruments range from 0 degrees in the TT phase to a maximum of 13 degrees in the GF phase. Considering solely the data coming from Xsens in the two environments the ROMs differences range between a minimum of 0 degrees in both TP and GF phases to a maximum of 20 degrees in the FP phase.

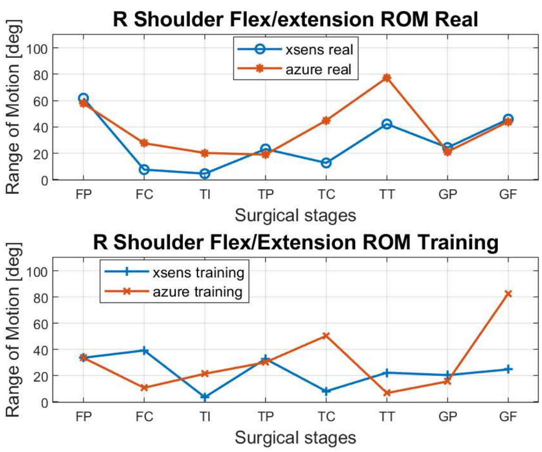

Fig. 4. Right shoulder flex/extension range of motion in both operating room (above) and simulation scenario (below). Data from Azure Kinect are represented in orange while data from Xsens in blue.

The differences between ROMs coming from Xsens and Azure in the right shoulder flex/extension in the operating room range between a minimum of 2 degrees in the GF phase to a maximum of 30 degrees in the TC phase, without considering the peak at around 80 degrees for the TT phase. Regarding the simulation environment, the differences in terms of ROMs between the two instruments range from 0 degrees in the FP phase to a maximum of 57 degrees in the GF phase. Considering solely the data coming from Xsens in the two environments the ROMs differences range between a minimum of 0 degrees in the TI phase to a maximum of 31 degrees in the FC phase.

2) Left shoulder range of motion

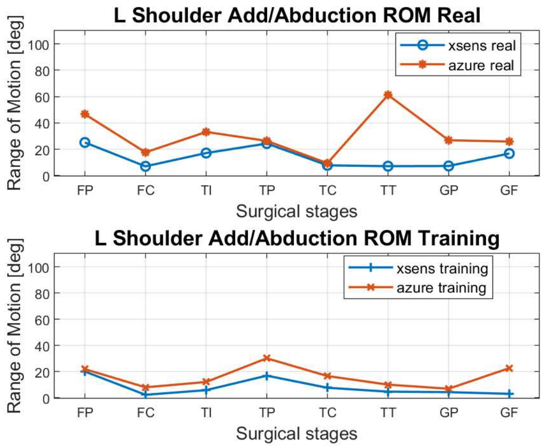

Fig. 5. Left shoulder add/abduction range of motion in both operating room (above) and simulation scenario (below). Data from Azure Kinect are represented in orange while data from Xsens in blue.

The differences between ROMs coming from Xsens and Azure in the left shoulder add/abduction in the operating room range between a minimum of 2 degrees in the TP phase to a maximum of 21 degrees in the FP phase, without considering the peak at 60 degrees for the TT phase. Regarding the simulation environment, the differences in terms of ROMs between the two instruments range from 3 degrees in the GP phase to a maximum of 19 degrees in the GF phase. Considering solely the data coming from Xsens in the two environments the ROMs differences range between a minimum of 0 degrees in the TC phase to a maximum of 13 degrees in the GF phase.

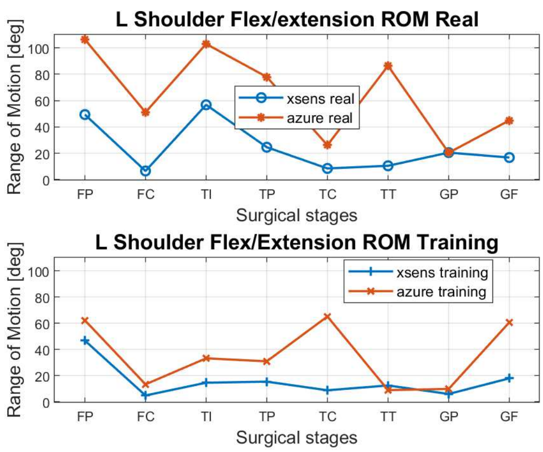

Fig. 6. Left shoulder flex/extension range of motion in both operating room (above) and simulation scenario (below). Data from Azure Kinect are represented in orange while data from Xsens in blue.

The differences between ROMs coming from Xsens and Azure in the left shoulder flex/extension in the operating room range between a minimum of around 0 degrees in the GP phase to a maximum of 56 degrees in the FP phase, without considering the peak at 85 degrees for the TT phase. Regarding the simulation environment, the differences in terms of ROMs between the two instruments range from 3 degrees in the TT phase to a maximum of 56 degrees in the TC phase. Considering solely the data coming from Xsens in the two environments the ROMs differences range between a minimum of 0 degrees in the TC phase to a maximum of 42 degrees in the TI phase.

3) Right elbow range of motion

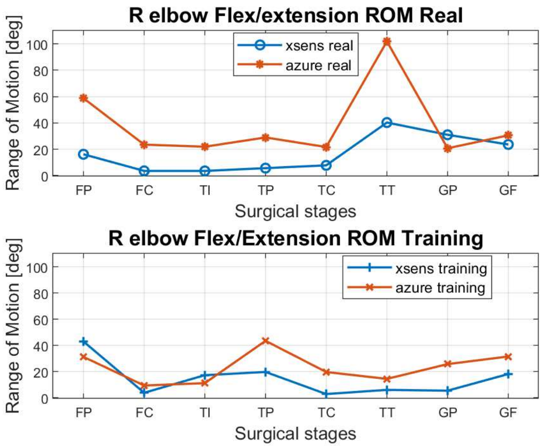

Fig. 7. Right elbow flex/extension range of motion in both operating room (above) and simulation scenario (below). Data from Azure Kinect are represented in orange while data from Xsens in blue.

The differences between ROMs coming from Xsens and Azure in the right elbow flex/extension in the operating room range between a minimum of 10 degrees in the GP phase to a maximum of 42 degrees in the FP phase, without considering the peak at 100 degrees for the TT phase. Regarding the simulation environment, the differences in terms of ROMs between the two instruments range from 5 degrees in the FC phase to a maximum of 23 degrees in the TP phase. Considering solely the data coming from Xsens in the two environments the ROMs differences range between a minimum of 0 degrees in the FC phase to a maximum of 34 degrees in the FP phase.

4) Left elbow range of motion

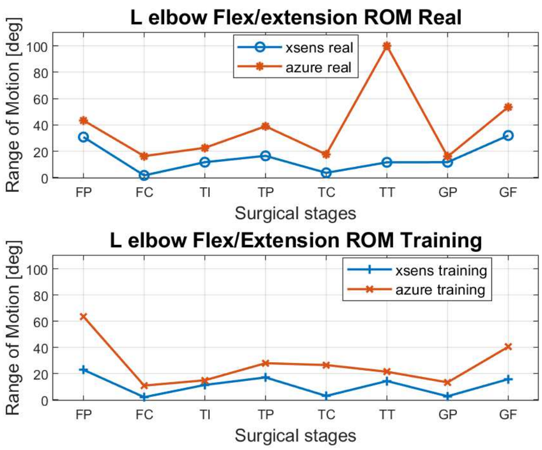

Fig. 8. Left elbow flex/extension range of motion in both operating room (above) and simulation scenario (below). Data from Azure Kinect are represented in orange while data from Xsens in blue.

The differences between ROMs coming from Xsens and Azure in the left elbow flex/extension in the operating room range between a minimum of 4 degrees in the GP phase to a maximum of 22 degrees in the TP phase, without considering the peak at 100 degrees for the TT phase. Regarding the simulation environment, the differences in terms of ROMs between the two instruments range from 3 degrees in the TI phase to a maximum of 40 degrees in the FP phase. Considering solely the data coming from Xsens in the two environments the ROMs differences range between a minimum of 0 degrees in diverse phases (from FC to TC) to a maximum of 16 degrees in the GF phase.

IV. DISCUSSION

The present work showed the stages needed to obtain kinematics measurements inside the operating room and to compare them to those obtained in the simulation environments. Starting from the definition of the access protocol and the interview with the surgeon to understand the feasibility of using IMUs and depth cameras, we managed to enter the operating room without obstructing the work of the surgical staff and obtaining reliable kinematic data. The collected data were then appropriately elaborated to obtain the results previously shown.

Considering the comparison between the two measurement systems, the primary observation from the graphs of the Azure Data in the operating room is the difference of almost 70 degrees with respect to Xsens data that occurs in all the shoulders and elbow movements in correspondence with the TT phase. This is due to occlusions due by the surgeon's assistants which were ahead of the surgeon in phase two. This is one of the main criticisms regarding the use of depth cameras. In fact, it must be positioned in a way that does not obstruct the work of the individuals in the room, while also providing the optimal perspective for observing the surgeon. Unfortunately, not always, the two requirements are compatible and since the first one cannot be disregarded, this results sometimes in obstruction to the camera, and some stages cannot be properly recorded. However, a depth camera has also an advantage with respect to a simple RGB camera. In fact, in such a critical environment as the operating room, there are a lot of people, all with the same clothes and all near each other. This implies that a markerless skeletonizer such as the MediaPipe BlazePose framework [14], which bases the video inference just on RGB images is not able to correctly segment the different people (see Figure 9).

Fig. 9. Instance of frame inference using MediaPipe BlazePose within the operating room. It's important to note that this framework cannot differentiate between two individuals wearing identical attire and overlapping each other. Consequently, this results in all the keypoints of the skeleton being wrongly positioned.

This problem is solved thanks to the depth camera, as already shown previously in Figure 3.

Another consideration is that the add/abduction of the shoulders showed less difference between the two instruments in terms of range of motion. This can be attributed to the Azure Kinect being positioned frontally to the surgeon, making abduction an easier and clearer movement to observe. That's a good result since as shown in the literature [11] the add/abduction is one of the main movements used to distinguish between surgeons with different levels of expertise. That hints that with some adjustments in positioning the Azure inside the operating room and improvements to the preliminary OpenSim model used, the Azure could be a possible solution to monitor surgeons' movements inside the operating room.

The final observation regards the discrepancy in ranges of motion between the simulation environment and the operating room considering data coming from Xsens. While most stages show minimal differences, certain stages exhibit discrepancies of up to 40 degrees. Notably, these differences predominantly pertain to the flexion/extension movements of the shoulder. As previously explained, the camera's positioning favors adduction/abduction movements, potentially contributing to this observed variation.

A limitation of the present study is that we have just a single test on the same surgeon for both environments and further studies required more trials to statistically assess the reliability of the data. Now that the protocol is defined, an acquisition campaign will start with the aim of acquiring data on different surgeons with different levels of experience to increase the variability of the dataset and obtain significant statistical results.

V. CONCLUSIONS

The present study highlighted the feasibility of entering the operating room during an arthroscopic procedure by following some recommendations summed up in the access protocol. The results demonstrated the feasibility of obtaining kinematic metrics through IMUs inside the operating room, despite some discrepancies with the training environment using a knee mannequin. Additionally, the potential use of an Azure Kinect DK as an alternative to IMUs was explored, revealing certain challenges. Overall, this study provides valuable insights into the practicality and challenges of implementing kinematic measurements in surgical environments. Of course, just one experiment is not enough, so further research is needed to obtain statistical data which would provide robust conclusions and validation of the proposed measurement method.

REFERENCES

[1] A. Cuschieri, N. Francis, J. Crosby, and G. B. Hanna, "Surgical education What do master surgeons think of surgical competence and revalidation?," 2001.

[2] R. K. Reznick and C. Ontario, "Teaching and Testing Technical Skills."

[3] K. Moorthy, Y. Munz, S. K. Sarker, and A. Darzi, "Objective assessment of technical skills in surgery," *British Medical Journal*, vol. 327, no. 7422. BMJ Publishing Group, pp. 1032–1037, Nov. 01, 2003. doi: 10.1136/bmj.327.7422.1032.

[4] "Assessing operative skill, needs to be more objective.".

[5] M. Bridges and D. L. Diamond, "The Financial Impact of Teaching Surgical Residents in the Operating Room," 1999.

[6] J. L. Hodgins and C. Veillette, "Arthroscopic proficiency: methods in evaluating competency," 2013. [Online]. Available: http://www.biomedcentral.com/1472-6920/13/61

[7] R. J. Koehler *et al.*, "The Arthroscopic Surgical Skill Evaluation Tool (ASSET)," *American Journal of Sports Medicine*, vol. 41, no. 6, pp. 1229–1237, Jun. 2013, doi: 10.1177/0363546513483535.

[8] S. D. Bann, M. S. Khan, and A. W. Darzi, "Measurement of surgical dexterity using motion analysis of simple bench tasks," *World J Surg*, vol. 27, no. 4, pp. 390–394, Apr. 2003, doi: 10.1007/s00268-002-6769-7.

[9] Y. Tashiro, H. Miura, Y. Nakanishi, K. Okazaki, and Y. Iwamoto, "Evaluation of skills in arthroscopic training based on trajectory and force data," *Clin Orthop Relat Res*, vol. 467, no. 2, pp. 546–552, 2009, doi: 10.1007/s11999-008-0497-8.

[10] G. S. J. Kirby *et al.*, "Assessing arthroscopic skills using wireless elbow-worn motion sensors," *Journal of Bone and Joint Surgery - American Volume*, vol. 97, no. 13, pp. 1119–1127, Sep. 2014, doi: 10.2106/JBJS.N.01043.

[11] M. Rose *et al.*, "Wearable Inertial Sensors Allow for Quantitative Assessment of Shoulder and Elbow Kinematics in a Cadaveric Knee Arthroscopy Model," *Arthroscopy - Journal of Arthroscopic and Related Surgery*, vol. 33, no. 12, pp. 2110–2116, Dec. 2017, doi: 10.1016/j.arthro.2017.06.042.

[12] A. Seth, M. Sherman, J. A. Reinbolt, and S. L. Delp, "OpenSim: A musculoskeletal modeling and simulation framework for in silico investigations and exchange," in *Procedia IUTAM*, Elsevier B.V., 2011, pp. 212–232. doi: 10.1016/j.piutam.2011.04.021.

[13] A. Rajagopal, C. L. Dembia, M. S. DeMers, D. D. Delp, J. L. Hicks, and S. L. Delp, "Full-Body Musculoskeletal Model for Muscle-Driven Simulation of Human Gait," *IEEE Trans Biomed Eng*, vol. 63, no. 10, pp. 2068–2079, Oct. 2016, doi: 10.1109/TBME.2016.2586891.

[14] C. Lugaresi *et al.*, "MediaPipe: A Framework for Building Perception Pipelines," Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.08172