



HAL
open science

Stand-up straight!: human pose estimation to evaluate postural skills during orthopedic surgery simulations

Tiphaine Casy, Alexandre Tronchot, Herve Thomazeau, Xavier Morandi,
Pierre Jannin, Arnaud Huauhmé

► To cite this version:

Tiphaine Casy, Alexandre Tronchot, Herve Thomazeau, Xavier Morandi, Pierre Jannin, et al.. Stand-up straight!: human pose estimation to evaluate postural skills during orthopedic surgery simulations. *International Journal of Computer Assisted Radiology and Surgery*, 2022, 18 (2), pp.279-288. 10.1007/s11548-022-02762-5 . hal-03826898

HAL Id: hal-03826898

<https://hal.science/hal-03826898>

Submitted on 3 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

“Stand-up straight!”: Human pose estimation to evaluate postural skills during orthopedic surgery simulations

Tiphaine Casy¹, Alexandre Tronchot^{1,2}, Hervé Thomazeau^{1,2}, Xavier Morandi^{1,3}, Pierre Jannin¹, Arnaud Huaultmé¹

Affiliations

- (1) Univ Rennes, Inserm, LTSI – UMR 1099, F-35000 Rennes, France
- (2) Orthopedics and Trauma Department, Rennes University Hospital, F-35000 Rennes, France
- (3) Neurosurgery Department, Rennes University Hospital, F-35000 Rennes, France

E-mail (orcid number)

tiphaine.casy@univ-rennes1.fr (0000-0002-6662-3900); alexandre.tronchot@chu-rennes.fr (0000-0002-3092-3282); herve.thomazeau@chu-rennes.fr (0000-0002-7668-2867); xavier.morandi@chu-rennes.fr (0000-0002-8031-7221); pierre.jannin@univ-rennes1.fr (0000-0002-7415-071X); arnaud.huaultme@univ-rennes1.fr (0000-0002-7844-5259)

Abstract

Purpose: Surgery simulators can be used to learn technical and non-technical skills and, to analyse posture. Ergonomic skill can be automatically detected with a Human Pose Estimation algorithm to help improve the surgeon's work quality. The objective of this study was to analyse the postural behaviour of surgeons and identify expertise-dependent movements. Our hypothesis was that hesitation and the occurrence of surgical instruments interfering with movement (defined as interfering movements) decrease with expertise.

Material and Methods: Sixty surgeons with three expertise levels (novice, intermediate, and expert) were recruited. During a training session using an arthroscopic simulator, each participant's movements were video-recorded with an RGB camera. A modified OpenPose algorithm was used to detect the surgeon's joints. The detection frequency of each joint in a specific area was visualized with a heatmap-like approach and used to calculate a mobility score.

Results: This analysis allowed quantifying surgical movements. Overall, the mean mobility score was 0.823, 0.816, and 0.820 for novice, intermediate and expert surgeons, respectively. The mobility score alone was not enough to identify postural behaviour differences. A visual analysis of each participants' movements highlighted expertise-dependent interfering movements.

Conclusion: Video-recording and analysis of surgeon's movements are a non-invasive approach to obtain quantitative and qualitative ergonomic information in order to provide feedback during training. Our findings suggest that the interfering movements do not decrease with expertise but differ in function of the surgeon's level.

Keywords: 2D human pose estimation, Orthopedic Surgery, Process Assessment, Simulation training

Statements and Declarations

Conflict of Interest

The authors declare that they have no conflict of interest.

Ethical approval

All procedures involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

This article does not contain patient data.

Acknowledgement

This work was possible thanks to the VIRTAMED Society and especially Mrs. Vitz.

1. Introduction

Simulation-based learning techniques have been used in different fields of medicine for many years [1]. One of the oldest simulation models is a clay nasal reconstruction model from ancient India that dates back to 600 BC [2]. Thanks to technological advances, medical simulators have improved in realism and number of simulated procedures [1, 2]. Laparoscopic surgery is the specialty that most requires of simulation-based training for many reasons including the reduction of tutoring time by experts and the development of technical skills [3–5]. Indeed, simulation has many advantages such as improved ethics, perfect reproducibility, and acquisition of technical skills without endangering the patient's health [2–4]. Concerning technical skills, simulators record the execution time, gesture accuracy, and correct use of the laparoscope [5–7]. These competencies have been shown to be insufficient, lacking non-technical skills aspects such as communication, stress management and gesture [7–9]. Technical and non-technical skills are indissociable, as shown by Fecso *et al.*, using statistical correlations [10]. Another important skill is ergonomic self-awareness, to avoid musculoskeletal pain caused by repetitive surgeries. The most studied ergonomic factors are the movement and posture of the surgeon's head, shoulder, and neck [11, 12].

Surgeon's posture can be detected using cameras and/or electromyographic sensors (EMGs) [13–16]. These technologies allow analysing the surgeon's ergonomic dynamics by estimating postural angles and muscle tension [17, 18] and by providing information to limit musculoskeletal pain in the lower back, neck and shoulders. Some studies have analysed the various parameters that influence the surgeon's posture with the aim of proposing solutions to reduce pain and improve the overall surgical ergonomic [16, 17, 19]. These methods are based on data collected using EMGs and/or Inertial Measurement Units (IMU) directly attached to the surgeon's body. The camera-based method described by Berguer *et al.* [19] uses reflective markers visible thanks to incandescent bulbs to capture the surgeon's posture. However, these devices may hinder the surgeon's movements, creating a possible bias, and are time-consuming to attach. Overall, the data obtained with these different devices are similar: movement frequency, angle calculation between body joints, and muscles activity.

Human pose estimation (HPE) algorithms have been developed and improved since the 80's and have been used to detect and monitor surgeons' posture detection [20–22]. This technology has been rapidly implemented in different domains such as surveillance [22], athletics, and medical applications [21]. The HPE algorithm was the Pictorial Structure, using color and regression algorithms applied on 2 dimensional (2D) pictures [23]. The democratization of deep learning has increased HPE accuracy and performance. For instance, the 'DeepPose' tool (2014) has with a HPE accuracy of 60% [24–26]. Since then many other algorithms have been developed to improve posture detection in different situations. Deep learning HPE methods have been used to address several challenges including multi-person detection and realtime pose estimation [27–29]. For multi-person detection, two major approaches exist: "top-down" and "bottom-up". In "top-down" approaches, first people are detected and separated and then the body joints are detected. "Bottom-up" approaches move from joints to people [27–31]. The most popular deep learning HPE tool is OpenPose because it allows multi-person pose detection, is the most documented method, and is an open-source algorithm [31, 32]. It can also be used for real-time posture detection [27, 28]. This "top-down" algorithm was trained to recognize different body joint dataset which is useful for adapting joint detection in function of the available images or videos [28, 31–33].

The purpose of the present work was to observe different postural behaviours and to correlate their frequency with surgical expertise. These postural behaviours include those arising from discomfort hesitation, and those due to the surgical tool inference with the desired movement (i.e. "interfering movement"). Our hypothesis was that the number of interfering movements was higher for "novice" than for "expert" surgeons. To test this hypothesis, we developed an original and innovative mobility score to analyse and compare the surgeons' movements during a simulated surgical intervention. We also created a comprehensive heatmap that illustrates the median mobility scores. This representation is helpful to focus the analysis on exercises with the most differences among surgeons during simulation-based learning.

2. Materials and Methods

To test our hypothesis, a cohort of surgeons with different level of expertise were monitored while using a virtual reality arthroscopic simulator using a convenient non-invasive human pose detection method.

2.1. Virtual Reality Arthroscopic simulation training protocol

Sixty surgeons were recruited over 9 months (November 2020 to July 2021) for a controlled prospective observational study:

- 36 “novices”: junior residents (one or two years post-graduation);
- 12 “intermediate”: < 100 arthroscopies as principal operator per year **or** < 5 years after the postgraduate diploma in orthopaedic surgery;
- 12 “experts”: > 100 arthroscopies as principal operator per year **and** > 5 years after the postgraduate diploma in orthopaedic surgery.

These criteria to define the expertise level were previously used in a similar study [34]. The cohort was multicentric and mainly composed of right-handed men. Women were under-represented in the “intermediate” and “experts” groups (Tab. 1). Informed consent was obtained from all participants. The Fundamentals of Arthroscopy Surgery (FAST) and KNEE modules from the VirtaMed AG (Zurich) ArthroS™ arthroscopy simulator were used for the study. The FAST module was used for basic exercises, and KNEE was used for more complex tasks [35]. A standardised experimental set-up consisting of nine different exercises was proposed to each surgeons to evaluate all the technical arthroscopic competences. These exercises included also a simulated arthroscopic meniscectomy that was performed three times consecutively at the end of the session [34]. These exercises were selected because they require precise surgical skills during real arthroscopic interventions. Each participant was recorded during the session using a frontal standard RGB camera (recording frequency of 30 Hertz and a resolution of 240 x 320 pixels). Recording was done using a Python script (version 3.7) on a dedicated laptop (16GB RAM, Intel core i7). Each recorded video was anonymized using a standard procedure to enable blind analysis.

Tab. 1 Description of the surgeon cohort

Expertise level	Number	Mean age (min/max)	Hospital localisation	Man/Woman	Right/Left handed
Novice	36	25.17 (24/29)	Angers, Brest, Nantes, Rennes, Tours	26/10	21/5
Intermediate	12	32.5 (30/39)	Angers, Lyon, Nantes, Rennes, Tours	11/1	12/0
Expert	12	56.72 (39/70)	Angers, Bordeaux, Brest, Lille, Nantes, Paris, Rennes	11/1	11/1

2.2. Detection of the surgeons' posture

A simplified version of OpenPose [33] was used to detect the participants' joints in the videos of their training session. This algorithm could be adapted using the detection confidence threshold as parameter to return the orthonormal coordinates of each joint¹. The detection confidence threshold was automatically optimised to detect a precise list of body joints. Indeed, the simulator module hides the participants' lower body part (Fig.1) that therefore cannot be detected by the algorithm in the input RGB images.



Fig. 1 Camera view of a surgeon using the KNEE module. The lower limbs and head are not visible or hidden

¹ <https://github.com/quanhua92/human-pose-estimation-opencv>

Despite the confidence threshold parameter optimization, the resulting detection was considered too noisy. To filter and smooth the signal, a Savitzky-Golay filter was applied on abscissa and ordinates values of each detected joint [36]. Two parameters were needed for this algorithm: the length of the sliding window and the polynomial degree that represents the data distribution. Both are explained in the following to subsections.

- Length of the window

This parameter is essential to obtain a filtering close to the original data without making up information. There is no strict mathematical rule for this and the sliding window length is highly data dependent. To ensure that data were not over-smoothed, the window length was constrained as follow:

$$\frac{1}{t} \sum_{t=0}^t \text{abs}(x(t) - x(t)') > \sqrt{\frac{\sigma_x^2}{2}} \quad (1)$$

where σ_x^2 is the standard deviation of the original data, and $x(t)$ and $x(t)'$ are the original and Savitzky-Golay values, respectively, at the time t .

- Polynomial degree

A third degree polynomial was found to best fit the original data after testing polynomial degrees from 2 to 5. Fig. 2 presents the denoising process with the Savitzky-Golay filter and these parameters.

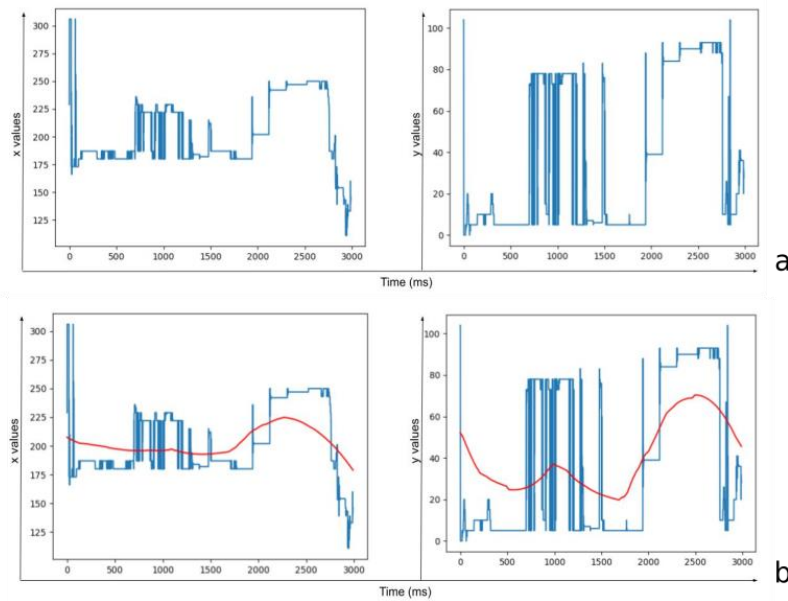


Fig. 2 Graphical representation of the camera coordinates of the neck-body joint of one surgeon during the “periscoping” exercise in function of time (in millisecond; ms). Original signal before (in blue) (a) and after denoising (in red) (b) using a Savitzky-Golay filter with automatic window size and a third degree polynomial

2.3. Analysis of the 2D coordinates

The final joint coordinates were visualised using a virtual skeleton (Fig. 3.a and a movie in Supplementary Material). From this representation an original approach was developed to visualise and analyse posture as a movement heatmap (Fig. 3.b) and to calculate a novel mobility score.

- Movement heatmap

The movement heatmap was created in two steps: i) partitioning of the camera coordinate space into a series of 10*10 pixel squares, and ii) computing the frequency at which each joint appeared in each squares. When a joint was included in one of the squares, its corresponding coefficient in the matrix was incremented by one. This step was repeated for each coordinate present in the exercise coordinate list (each time step). At the end of the process, the matrix is represented by a grey-scale picture, from light (low frequency) to dark (high frequency). This representation provides a global view of the movements performed by a surgeon during simulator exercises (Fig. 3.b).

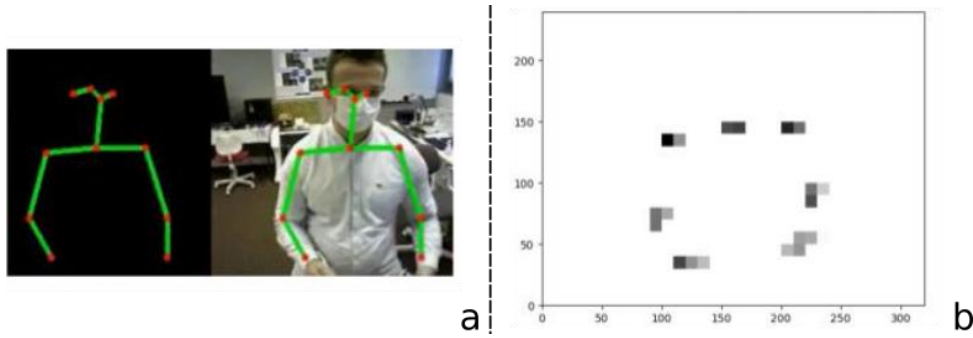


Fig. 3 Representation of the surgeon's pose estimation by the simplified OpenPose algorithm (a); movement heatmap linked to the surgeon posture (b)

- Mobility score

The movement heatmap showed the physical extent of joint motion in relation to the global environment. It was used to compute a mobility score (MS) that takes into account the theoretical maximum (\max_{theory}) value of the coefficient (when all joint coordinates are in the same 10×10 pixel square during the entire exercise) and the observed reached maximum (\max_{obs}), defined by the maximum coefficient value (i.e. the one with the most joint coordinates during the whole exercise). The third quartile (Q3) takes into account the coefficient value distributions. The MS was calculated with the following formula:

$$MS = 1 - \frac{|\max_{theory} - (\max_{obs} - Q3) \times 4|}{frames - squares} \quad (2)$$

where frames and squares are the total number of frames that constitute the exercise video and the total number of incremented coefficients, respectively.

The MS represents the distribution of joint coordinates in the global environment and ranges between 0 and 1. A score close to 0 indicates few movements, whereas values, close to 1 indicate a large number of joint movement. The MS was calculated for each surgeon, exercise, and joint combination. Importantly, the MS is height and camera position independent.

2.4. Statistics

As the MS did not follow a normal distribution (significant Shapiro-Wilk test), the Mann Whitney test was used to compare MS values among the expertise groups. Differences were considered significant if the p -value was < 0.05 .

3. Results

To facilitate comparison and highlight differences in the three expertise groups, the median MS values of each implicated joint in the different exercises were represented using a grey scale intensity (Fig. 4). Several differences can be seen between the three groups. For instance, the median MS values for the left shoulder of "Novice 1" (upper part of Fig.4) indicated that this surgeon moved a lot during the simulation exercises (MS between 0.9 and 1.0 for more than 80% of exercises (Fig. 4)). This movement profile was close to that of the "novices" group, but with higher scores for all exercises, with the exception of "image centering" and "probing".

To obtain an overview of the three expertise groups for all exercises, the mean MS for each group was calculated by grouping all joints and exercises: 0.823 (standard deviation: 0.22), 0.816 (standard deviation: 0.22), and 0.820 (standard deviation: 0.25) for the "novice", "intermediate" and "expert" groups, respectively. No significant difference was observed among these three mean global MS ($p > 0.05$). When focusing on the individual joints, wrists and left elbow were moved frequently during each exercise. Moreover, the MS profiles for the same exercise (e.g. "telescoping" and "periscoping") were different in the three expertise. Overall, the MS profiles of the three expertise groups suggested that "novice" and "expert" surgeons had the median in the same gradient for many exercises compared with "intermediate" moving globally less than the others.

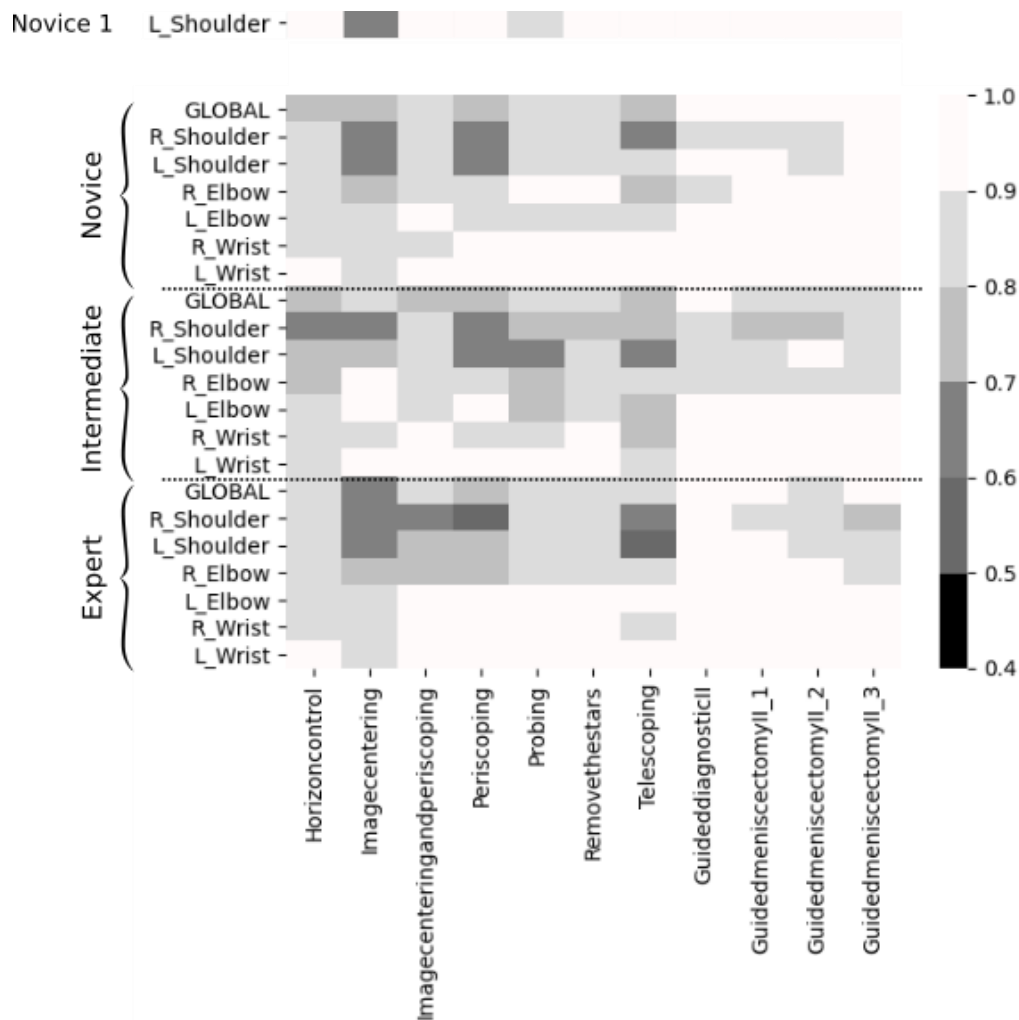


Fig. 4 Upper part: representation of the left shoulder score for one “novice” surgeon for all exercises. Lower part, “mobility heatmap”: median mobility score for each exercise and joint in the three expertise groups

On the basis of the MS profile, the next analysis focused on a subset of joint-exercise pairs where the MS differences among expertise levels were most pronounced. This subset included exercises that required the most technical skills and therefore are more relevant for surgical training. This subset included both shoulders for “periscoping”, “guided diagnostic” and “guided meniscectomy 1” and the left shoulder for “probing”. Fig. 5 shows boxplots of these scores, allowing for deeper visual analysis. Differences between expertise groups were most visible for “periscoping” where the “expert” group having significantly lower right shoulder mobility ($p = 0.0215$, $p = 0.0379$ compared with the novice and intermediate group, respectively) (Tab. 2). Similarly, the MS for left shoulder in the “guided diagnostic” exercise ($p = 0.0392$, $p = 0.0193$) and “guided meniscectomy 1” for the left shoulder ($p = 0.0174$, $p = 0.0113$) were higher in the expert group than the novice and intermediate groups. This analysis also emphasised the MS heterogeneity within the “expert” and “novice” groups, particularly for the “probing” exercise, as indicated by the large number of outliers.

Comparison of the median MS for each examined joint-exercise pair with the Mann Whitney test highlighted at least one significant difference between expertise groups. Overall, the “novice” and “intermediate” groups showed similar performance resulting in only one significant differences, for the “probing” exercise.

Tab. 2 Comparison of median MS in function of the expertise level with the Mann-Whitney test. Significant differences (p -value < 0.05) are highlighted in grey

Exercise	Body joint	Novice VS Intermediate	Novice VS Expert	Intermediate VS Expert
“Periscoping”	Right Shoulder	0.2672	0.0215	0.0379
	Left Shoulder	0.4429	0.1024	0.1473
“Guided diagnostic”	Right Shoulder	0.3164	0.0392	0.0193
	Left Shoulder	0.0854	0.2293	0.0538
“Guided meniscectomy 1”	Right Shoulder	0.3657	0.4659	0.3721
	Left Shoulder	0.3708	0.0174	0.0113
“Probing”	Left Shoulder	0.0288	0.2677	0.1551

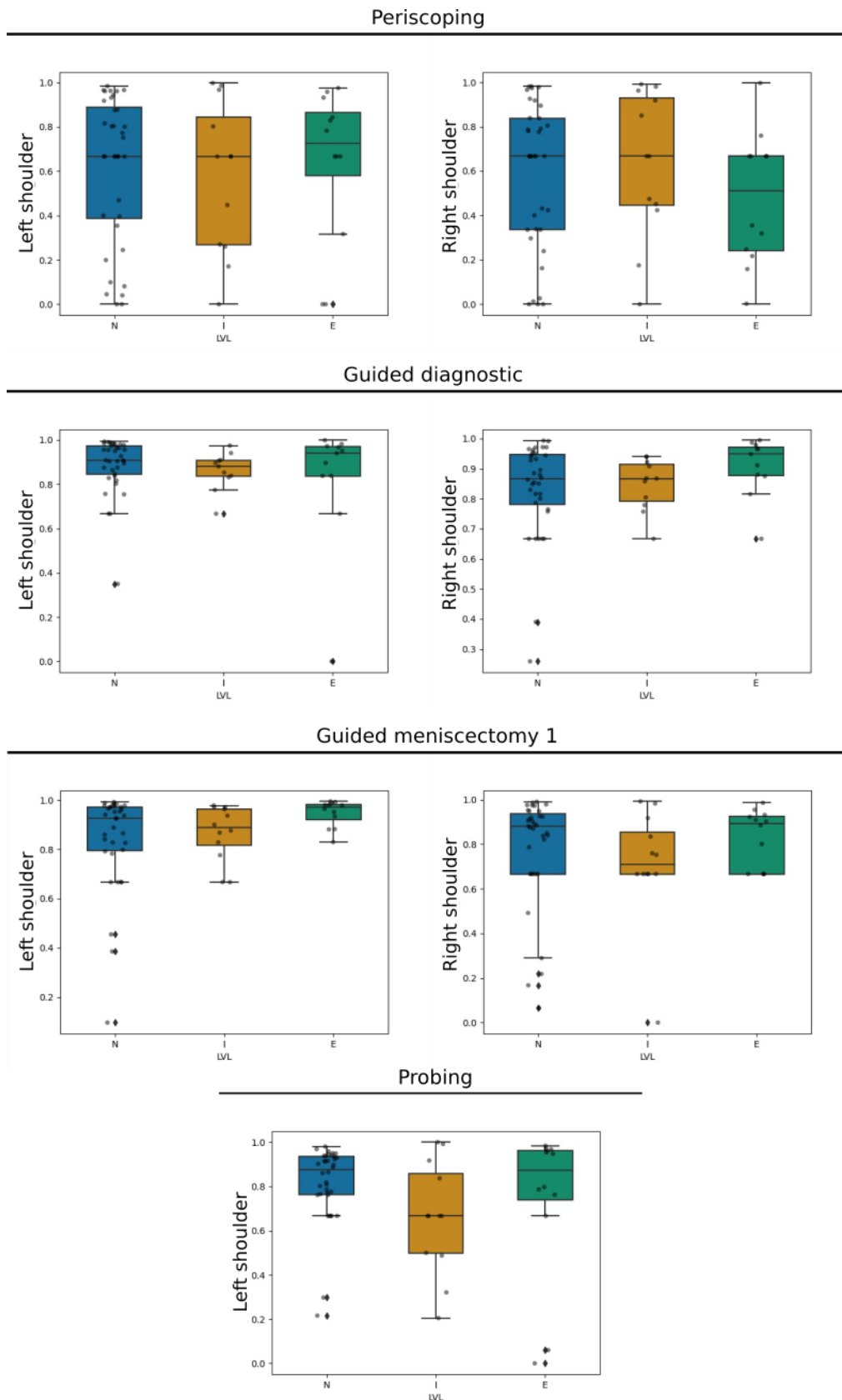


Fig. 5 Boxplots of the right and left shoulder MS in the “periscoping”, “guided diagnostic” and “guided meniscectomy 1” exercises and of the left shoulder MS for “probing”. Blue, orange, and green boxes correspond to the “novice” (N), “intermediate” (I), and “expert” (E) groups respectively

To qualitatively understand these differences among expertise groups the videos were visually analysed to detect interfering movements. "Novice" surgeons had interfering movements and hesitation movements close to their chest that sometimes interfered with each other leading to non-smooth displacements. The "intermediate" expertise level group was the most heterogeneous. Compared with the "novice" group the frequency and amplitude of interfering movements were decreased in the "intermediate" group, and overall their movements were more similar to those of expert surgeons. In the "expert" group, interfering movements were rarer but participants often physically or verbally interacted with the work environment or the study operator (i.e. scratched themselves, pointed to the simulator screen). Our clinical partners were not surprised by these observations especially for the "novice" group, and confirmed they were common in the operating room.

4. Discussion

This study wanted to assess surgical postural skills in line with other studies in this domain. Previous studies analysed surgical ergonomics to estimate risks [16, 17, 37], or to compare open and laparoscopic surgical procedures [19]. The most common methods rely on physical markers or sensors directly placed on the surgeon, such as Inertial Measurement Units (IMU), electromyographic sensors, or reflective bulbs [16, 17, 19, 37], that can be cumbersome. Concerning marker-less assessment, some studies used OpenPose to assess the surgeon's cognitive skills [38] or to perform automatic surgical procedure recognition [39]. Conversely, no study used marker-less methods to assess the surgeon's postural skills by comparing postural behaviour variability in function of the expertise level. Our study employed OpenPose and a novel quantification method to measure postural behaviours of different automatically detected joints in different surgical exercises. This allowed us to compare surgeons in a height-agnostic manner and independently of the camera positioning. Our study included several strengths such as: its multicentric data collection (sixty surgeons across nine French national hospitals) accessibility due to inexpensive and lighting recording equipment; and procedural reproducibility allowed by the simulator.

Shoulders showed the most significant differences among groups, whereas wrists displayed the highest MS for most exercises. This was explained by the impact of joint movements on the other joints. Indeed, when the shoulder had an accelerated movement, the impact on the elbow was amplified and even more on the wrist. The same scheme was observed when the elbow had an accelerated movement. This information explained why the biggest MS differences among groups and exercises concerned the shoulders and not the elbows and even less the wrists.

Mobility scores comparison highlighted significant differences between the "expert" group and the "novice" or "intermediate" groups (Tab. 2.) but not for all exercise-joint pair. This could be explained by the presence of outliers with extreme values within each expertise group, particularly for the right shoulder in the "novice" group during the periscoping exercise. The presence of outliers, above and below the median indicates that MS comparison was not fully adapted to study the postural skills and that this metric alone does not sufficiently describe interfering movements. In future this analysis will be paired with a gold standard ergonomic assessment because even expert clinicians often lack postural self-awareness skills, thus increasing the risk of injury [16, 17]. This was noted by our clinical collaborators especially concerning ergonomic differences resulting in using the endoscope and surgical instruments with both the dominant and non-dominant hands depending on the surgical side (i.e. left vs right knee).

Our qualitative results (interfering movements) can be explained by the different levels of knowledge and technical skills depending on the surgeon's expertise level, as shown by Morineau et al. [40]. They found that "intermediate" surgeons activated more conflicts and controls than "expert" and "novice" surgeons and this explained the reduced work amplitude for "intermediate" surgeons. Controls and conflicts were slightly lower in the "expert" than in the "intermediate" group and simulation exercises were less taxing than real surgery [40]. Interfering movements could also be explained by unintentional "gamification" that might have biased the last two repetitions of the "guided meniscectomy" exercise [41]. Indeed, surgeons wanted to increase the score returned by the simulator (by reducing the operation time) without worrying about their posture and surgical movements. More precise sub-group analysis will be necessary to understand these cognitive aspects that could allow defining a quantitative link between cognitive state and postural behaviour. Moreover, the combination of cognitive and postural domains could redefine the expertise groups and may represent another approach to understand surgical posture, which we intend to investigate in the future.

Our current study has several limitations. Specifically, information was lacking about the lower limbs due to the simulator and camera position. Moreover, the absence of depth information from the monoscopic camera set-up led to reduced appreciation of three-dimensional movements [34,35]. The OpenPose joint detection could also be

improved technically. This might lead to the removal of the parameterized filtering algorithm, and the increase of the overall method robustness. Lastly, it would be more pertinent to do the HPE in real-time to provide educational feedback to trainees. However, our 2D based method has many financial and utilisation advantages allowing a future leading to potentially more widespread use in the surgical theatre without interfering with the surgery itself. This would allow its use not only for training but also for continuous ergonomic analysis.

5. Conclusion

The study hypothesis was partially verified. Interfering movements were observed also in surgeons with higher expertise level, but they were different in the three groups. Exercises have been carefully selected, examining those that require precise technical skills reflective of real arthroscopic interventions and display the most differences observed using the global heatmap (Fig. 4).. The chosen 2D method with a standard camera was a non-invasive way to detect the surgeons' postures and therefore, may be used directly in operating room. The creation of movement heatmap and the corresponding MS provided interesting information about the physical skills of arthroscopic surgeons. Our approach highlighted several qualitative and quantitative differences in postural behaviour that reflect the surgical expertise level. In the future, this approach may be used for training purposes, and also for measuring and improving surgical ergonomics in the operating room.

6. Bibliography

1. Satava RM (2008) Historical Review of Surgical Simulation—A Personal Perspective. *World J Surg* 32:141–148. <https://doi.org/10.1007/s00268-007-9374-y>
2. Badash I, Burt K, Solorzano CA, Carey JN (2016) Innovations in surgery simulation: a review of past, current and future techniques. *Ann Transl Med* 4:. <https://doi.org/10.21037/atm.2016.12.24>
3. Sutherland LM, Middleton PF, Anthony A, et al (2006) Surgical Simulation. *Ann Surg* 243:291–300. <https://doi.org/10.1097/01.sla.0000200839.93965.26>
4. Satava RM (2001) Surgical education and surgical simulation. *World J Surg* 25:1484–1489. <https://doi.org/10.1007/s00268-001-0134-0>
5. Brewin J, Tang J, Dasgupta P, et al (2015) Full immersion simulation: validation of a distributed simulation environment for technical and non-technical skills training in Urology. *BJU Int* 116:156–162. <https://doi.org/10.1111/bju.12875>
6. Narazaki K, Oleynikov D, Stergiou N (2006) Robotic surgery training and performance. *Surg Endosc Interv Tech* 20:96–103. <https://doi.org/10.1007/s00464-005-3011-3>
7. Smith CD, Farrell TM, McNatt SS, Metreveli RE (2001) Assessing laparoscopic manipulative skills. *Am J Surg* 181:547–550. [https://doi.org/10.1016/s0002-9610\(01\)00639-0](https://doi.org/10.1016/s0002-9610(01)00639-0)
8. Agha RA, Fowler AJ, Sevdalis N (2015) The role of non-technical skills in surgery. *Ann Med Surg* 4:422–427. <https://doi.org/10.1016/j.amsu.2015.10.006>
9. Sharma B, Mishra A, Aggarwal R, Grantcharov TP (2011) Non-technical skills assessment in surgery. *Surg Oncol* 20:169–177. <https://doi.org/10.1016/j.suronc.2010.10.001>
10. Fecso AB, Kuzulugil SS, Babaoglu C, et al (2018) Relationship between intraoperative non-technical performance and technical events in bariatric surgery. *Br J Surg* 105:1044–1050. <https://doi.org/10.1002/bjs.10811>
11. Azari DP, Miller BL, Le BV, et al (2020) Quantifying surgeon maneuvers across experience levels through marker-less hand motion kinematics of simulated surgical tasks. *Appl Ergon* 87:103136. <https://doi.org/10.1016/j.apergo.2020.103136>
12. Meltzer AJ, Hallbeck MS, Morrow MM, et al (2020) Measuring Ergonomic Risk in Operating Surgeons by Using Wearable Technology. *JAMA Surg* 155:444–446. <https://doi.org/10.1001/jamasurg.2019.6384>
13. Farella E, Pieracci A, Benini L, Acquaviva A (2006) A Wireless Body Area Sensor Network for Posture Detection. In: 11th IEEE Symposium on Computers and Communications (ISCC'06). pp 454–459
14. Xu Y, Chen J, Yang Q, Guo (2019) Human Posture Recognition and fall detection Using Kinect V2 Camera | IEEE Conference Publication | IEEE Xplore.
15. Nisky I, Hsieh MH, Okamura AM (2013) A framework for analysis of surgeon arm posture variability in robot-assisted surgery | IEEE Conference Publication | IEEE Xplore.
16. Dalager T, Jensen PT, Eriksen JR, et al (2020) Surgeons' posture and muscle strain during laparoscopic and robotic surgery. *Br J Surg* 107:756–766. <https://doi.org/10.1002/bjs.11394>
17. Athanasiadis DI, Monfared S, Asadi H, et al (2021) An analysis of the ergonomic risk of surgical trainees and experienced surgeons during laparoscopic procedures. *Surgery* 169:496–501. <https://doi.org/10.1016/j.surg.2020.10.027>

18. Lobo D, Anuarbe P, López-Higuera JM, et al (2019) Estimation of surgeons' ergonomic dynamics with a structured light system during endoscopic surgery. *Int Forum Allergy Rhinol* 9:857–864. <https://doi.org/10.1002/alr.22353>
19. Berguer R, Rab GT, Abu-Ghaida H, et al (1997) A comparison of surgeons' posture during laparoscopic and open surgical procedures. *Surg Endosc* 11:139–142. <https://doi.org/10.1007/s004649900316>
20. Sarafianos N, Boteanu B, Ionescu B, Kakadiaris IA (2016) 3D Human pose estimation: A review of the literature and analysis of covariates. *Comput Vis Image Underst* 152:1–20. <https://doi.org/10.1016/j.cviu.2016.09.002>
21. Aggarwal JK, Cai Q (1997) Human motion analysis: a review. In: *Proceedings IEEE Nonrigid and Articulated Motion Workshop*. pp 90–102
22. Gavrilu DM (1999) The Visual Analysis of Human Movement: A Survey. *Comput Vis Image Underst* 73:82–98. <https://doi.org/10.1006/cviu.1998.0716>
23. Fischler MA, Elschlager RA (1973) The Representation and Matching of Pictorial Structures. *IEEE Trans Comput C-22*:67–92. <https://doi.org/10.1109/T-C.1973.223602>
24. Khan NU, Wan W (2018) A Review of Human Pose Estimation from Single Image. In: *2018 International Conference on Audio, Language and Image Processing (ICALIP)*. pp 230–236
25. Liu Y, Xu Y, Li S (2018) 2-D Human Pose Estimation from Images Based on Deep Learning: A Review. In: *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. pp 462–465
26. Toshev A, Szegedy C (2014) DeepPose: Human Pose Estimation via Deep Neural Networks. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Columbus, OH, USA, pp 1653–1660
27. Cao Z, Simon T, Wei S-E, Sheikh Y (2017) Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Honolulu, HI, pp 1302–1310
28. Cao Z, Hidalgo G, Simon T, et al (2021) OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans Pattern Anal Mach Intell* 43:172–186. <https://doi.org/10.1109/TPAMI.2019.2929257>
29. Chen Y, Tian Y, He M (2020) Monocular Human Pose Estimation: A Survey of Deep Learning-based Methods. *Comput Vis Image Underst* 192:102897. <https://doi.org/10.1016/j.cviu.2019.102897>
30. Dang Q, Yin J, Wang B, Zheng W (2019) Deep learning based 2D human pose estimation: A survey. *Tsinghua Sci Technol* 24:663–676. <https://doi.org/10.26599/TST.2018.9010100>
31. Tanugraha P (2019) Understanding OpenPose (with code reference)— Part 1. In: *Anal. Vidhya*. <https://medium.com/analytics-vidhya/understanding-openpose-with-code-reference-part-1-b515ba0bbc73>. Accessed 8 Oct 2021
32. Lichtman A (2020) An Overview of Human Pose Estimation with Deep Learning. In: *BeyondMinds*. <https://beyondminds.ai/blog/an-overview-of-human-pose-estimation-with-deep-learning/>. Accessed 8 Oct 2021
33. Viswakumar A, Rajagopalan V, Ray T, Parimi C (2019) Human Gait Analysis Using OpenPose. In: *2019 Fifth International Conference on Image Information Processing (ICIIP)*. pp 310–314
34. Tronhot A, Berthelemy J, Thomazeau H, et al (2021) Validation of virtual reality arthroscopy simulator relevance in characterising experienced surgeons. *Orthop Traumatol Surg Res* 103079. <https://doi.org/10.1016/j.otsr.2021.103079>
35. Anetzberger H, Reppenhagen S, Eickhoff H, et al (2021) Ten hours of simulator training in arthroscopy are insufficient to reach the target level based on the Diagnostic Arthroscopic Skill Score. *Knee Surg Sports Traumatol Arthrosc Off J ESSKA*. <https://doi.org/10.1007/s00167-021-06648-y>
36. Sadeghi M, Behnia F (2018) Optimum window length of Savitzky-Golay filters with arbitrary order. <https://doi.org/10.48550/arXiv.1808.10489>
37. Carbonaro N, Mascherini G, Bartolini I, et al (2021) A Wearable Sensor-Based Platform for Surgeon Posture Monitoring: A Tool to Prevent Musculoskeletal Disorders. *Int J Environ Res Public Health* 18:3734. <https://doi.org/10.3390/ijerph18073734>
38. Dias R, Yule S, Kennedy-Metz L, Zenati M (2019) Psychophysiological Data and Computer Vision to Assess Cognitive Load and Team Dynamics in Cardiac Surgery. <https://www.semanticscholar.org/paper/Psychophysiological-Data-and-Computer-Vision-to-and-Dias-Yule/43d6d81bc5b78f5a222e93c719f9348e5f2ce564>. Accessed 5 Jul 2022
39. Kishi S, Suzuki N, Tsuyuki S, et al (2021) Convolutional Neural Network based on Temporal Pose Features for Surgical Procedure Recognition. *Proc ISCIE Int Symp Stoch Syst Theory Its Appl* 2021:60–64. <https://doi.org/10.5687/sss.2021.60>
40. Morineau T, Morandi X, Le Moëllic N, et al (2009) Decision Making During Preoperative Surgical Planning. *Hum Factors* 51:67–77. <https://doi.org/10.1177/0018720809332847>

41. Kapralos B, Moussa F, Dubrowski A (2014) An Overview of Virtual Simulation and Serious Gaming for Surgical Education and Training. In: Brooks AL, Brahnam S, Jain LC (eds) *Technologies of Inclusive Well-Being: Serious Games, Alternative Realities, and Play Therapy*. Springer, Berlin, Heidelberg, pp 289–306