



# Inter- and intra-operator variability of the segmentation of the knee joint

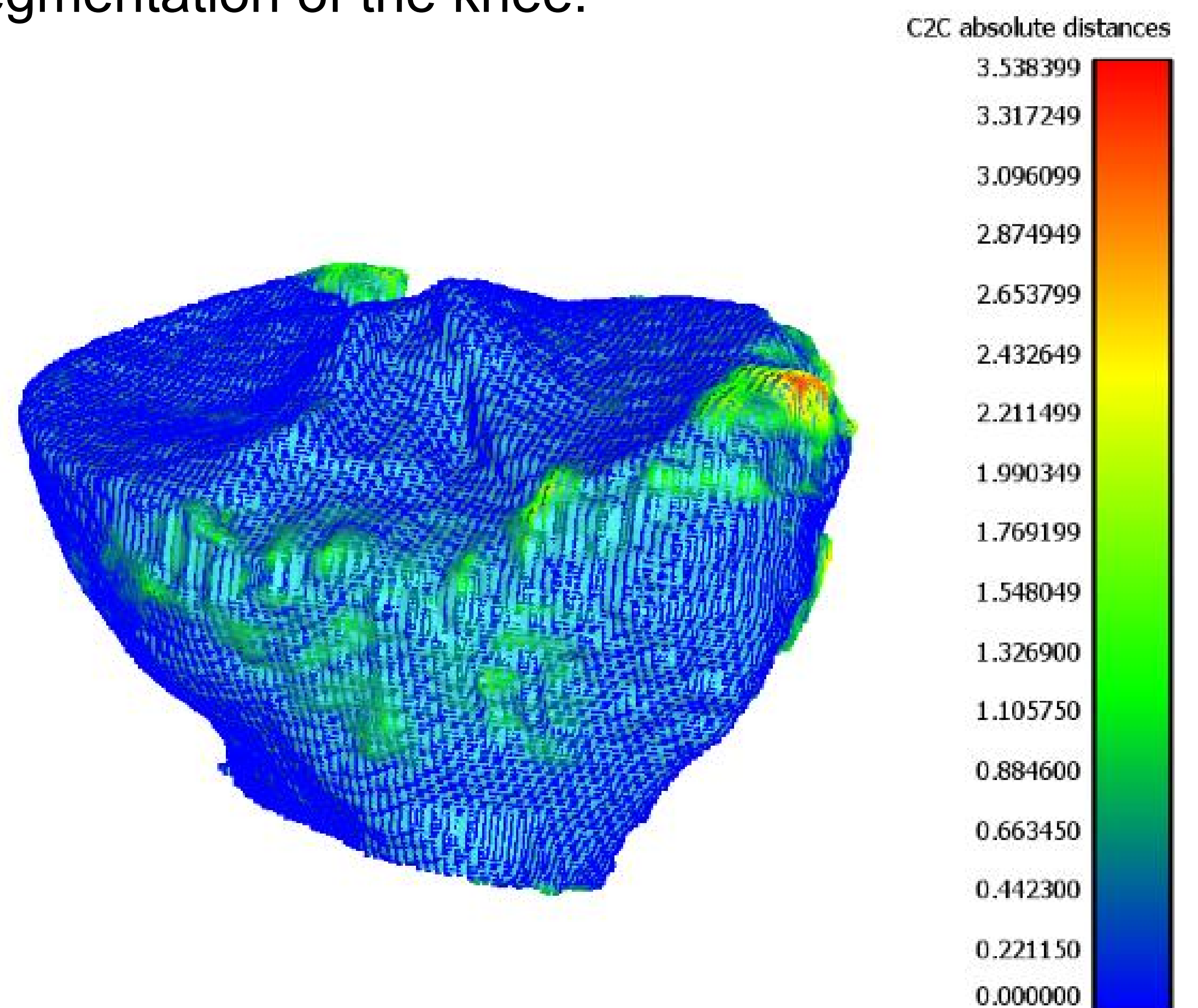
R. Asadirad<sup>1</sup>, D. Kahani<sup>1</sup>, P. Riches<sup>1</sup>

<sup>1</sup> University of Strathclyde, Wolfson centre, Glasgow, G4 0NW, UK

**Introduction:** Computer-Assisted Orthopaedic Surgery (CAOS) has increased significantly in the last decade with many CAOS procedures relying on pre-operative imaging. Since segmentation is semi-automatic, outcomes may vary [1] with possible implications for surgical planning, this study aimed to quantify the intra- and inter-operator variability of image segmentation of the knee.

## Methods

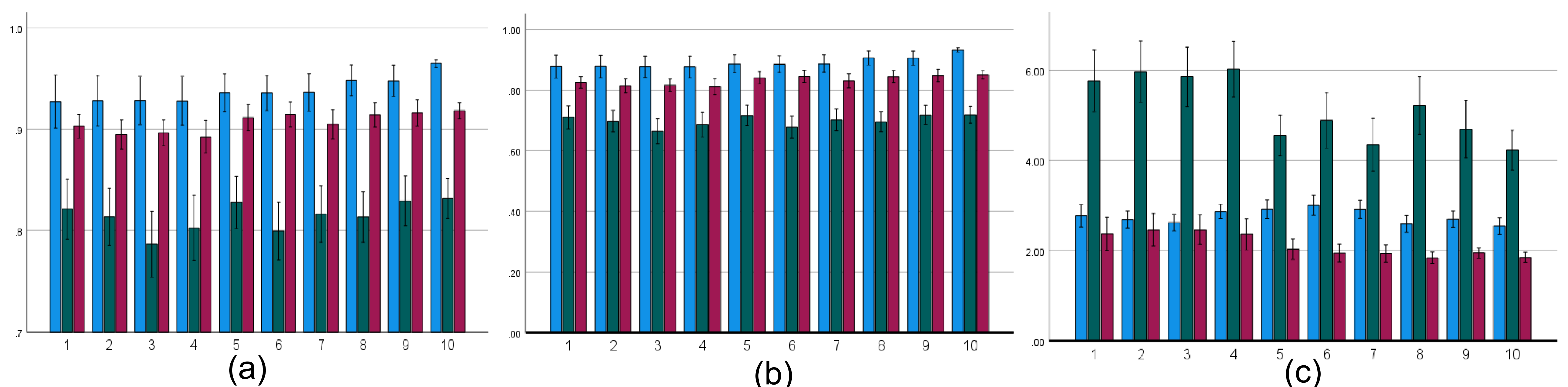
Following training, 15 volunteers segmented a single CT image (femur, tibia and patella), using the active contour and threshold method in ITK-Snap, five times. After each segmentation, the model was further manually edited. The Euclidian distance between point clouds was determined using CloudCompare™ (Figure 1). Dice Similarity Coefficient (DSC), Jaccard Index (JI) [2], and Hausdorff Distance (HD) were computed between all combinations of the five segmentations (10 comparisons). Repeated-measures ANOVA determined within-operator differences and interclass correlation coefficients (ICC) assessed inter-operator reliability.



**Figure 1:** Euclidian distances between point clouds for a representative comparison of tibial topology

## Results

DSC, JI, and HD varied with the bone being segmented (all  $p < 0.001$ ), whilst DSC and JI did not vary with comparison, but HD did (DSC:  $p = 0.08$ ; JI:  $p = 0.07$ ; HD = 0.03) (Figure 2a,b,c). HD potentially demonstrated self-learning by decreasing with later comparisons (Figure 2c). Intra-class correlation coefficients, display modest values for the femur, 0.456, tibia, 0.419 and patella, 0.451.



**Figure 2:** Mean (a) DSC (b) JI and (c) HD differences between the bones (blue: femur; green: tibia; red: patella) across the 10 comparisons

## Discussion & Conclusion

Mean differences between segmentations within an experienced operator were below the size of the voxel (1.2mm), although could reach ~5mm in isolated regions (Figures 1 & 2c). Intra-class correlations between operators suggest that the operator may play a crucial role in determining the outcome of the segmentation. The effect of such variation in segmentation in surgical planning or further biomechanical analyses (e.g FE) is unknown and should therefore be quantified.

## References

1. Gitto, S., et al., *Effects of Interobserver Variability on 2D and 3D CT- and MRI-Based Texture Feature Reproducibility of Cartilaginous Bone Tumors*. J Digit Imaging, 2021. 34(4): p. 820-832
2. Bertels, J. et al. (2019). Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice. In: Shen, D., et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019. Lecture Notes in Computer Science, vol 11765. Springer, Cham. [https://doi.org/10.1007/978-3-030-32245-8\\_11](https://doi.org/10.1007/978-3-030-32245-8_11).