

FROM LANDMARK TO IMPLANT: A SENSITIVITY ANALYSIS FRAMEWORK EXPOSING A HIDDEN FRAGILITY OF CT-BASED TKA PLANNING

Y. Roch¹³, D. Lozano³, N. Fougeron⁵, B. Gaulin¹², C. Horteur¹², Y. Payan¹, M. Bucki³, A. Perrier¹³⁴

¹ Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, Grenoble, France ² Department of Orthopedics and Sport Surgery, CHU Grenoble Alpes, Échirolles, France ³ Twinsight, Grenoble, France ⁴ Department of Orthopedics and Traumatology, Groupe Hospitalier Diaconesses Croix Saint-Simon, Paris, France ⁵ Univ. Rennes, Inria, CNRS, IRISA-UMR 6074, Rennes, France

Contact: yse.roch@univ-grenoble-alpes.fr

Introduction

Personalized Total Knee Arthroplasty (TKA) planning relies on a chain of computational steps linking CT-derived anatomical landmarks to morpho-functional metrics and ultimately to implant positioning decisions. Each step in this chain propagates uncertainty, yet current validation practices evaluate landmark accuracy in isolation, applying uniform distance thresholds with no consideration of downstream clinical impact. While foundational sensitivity analyses have established that short lever arms amplify positioning errors geometrically [1] and that reference frame landmarks exert disproportionate influence on system output [2], no framework exists that quantifies individual landmark-axis contributions across the full metric spectrum nor translates these findings into differentiated acceptance criteria. This study presents a novel methodological framework developed to quantify how landmark positioning errors propagate differentially into 14 surgical planning metrics across 30 patients and to establish an evidence-based criticality hierarchy directly applicable to quality control protocols for automated planning systems.

Methods

The framework integrates three methodological contributions. First, an atlas-based inverse morphing pipeline enables cross-patient aggregation of landmark positioning error distributions, preserving individual anatomical landmark variability while enabling population-level sensitivity quantifications. Second, a Monte Carlo uncertainty propagation engine (50,000 iterations per metric) propagates patient-specific empirical error distributions, derived from annotations by two orthopedic surgeons (repeatability: 1.83 mm; reproducibility : 2.43 mm), across 34 landmarks and 14 morpho-functional metrics spanning coronal, sagittal, and axial planes. Third, two complementary sensitivity indices characterize mechanistically distinct propagation pathways: an amplitude-based index (deg/mm, mm/mm) quantifies geometric lever-arm effects, while the Chatterjee coefficient (ξ) detects nonlinear reference frame contamination effects, invisible to variance-based methods (Fig. 1). Novel visualization tools (newly introduced Hanabi plots, sensitivity heatmaps, and dual-index criticality maps) translate quantitative findings into clinically interpretable representations (Fig. 3).

Results

The framework revealed systematic doubling between landmark positioning accuracy and metric reliability across all 14 metrics (CI₉₀ range: 1.37°–16.87° for angular metrics (Fig. 2); 1.09–3.31 mm for distance metrics). Two mechanistically distinct propagation pathways were identified and quantified: geometric amplification driven by short lever-arm axes (amplitude index up to 1.395 deg/mm for malleolar landmarks, Fig. 3) and reference frame contamination driven by plane-defining landmarks (Chatterjee ξ up to 0.356 for the lateral condyle Y-axis despite an amplitude of only 0.053 deg/mm), which simultaneously biases multiple geometrically independent metrics (Fig. 1). The dual-index framework demonstrated that landmarks with the highest positioning impact are not those with the largest raw positioning errors. It is a finding with direct implications for how automated planning tools should be validated prior to clinical deployment

Conclusion

This work introduces a transferable analysis framework (combining atlas-based morphing, Monte Carlo propagation, and dual-index characterization) applicable beyond TKA to any landmark-dependent surgical planning context (hip arthroplasty, spinal surgery, and craniofacial reconstruction). The framework exposes a structural flaw in current validation paradigms: uniform accuracy benchmarks are blind to the differential clinical impact of landmark errors. At the intersection of academic research, clinical data acquisition, and medical device development, this work illustrates how translational positioning, simultaneously accountable to surgical relevance, algorithmic performance, and regulatory standards, enables methodological contributions that neither purely academic nor purely industrial contexts would generate alone.

References

- [1] Luo, Tianyi & Martensson, Nicole & Howard, James & Stevens, Delaney & McIsaac, Kenneth & Lanting, Brent (2025). Identifying Sources of Error in Computer-Navigated Total Knee Arthroplasty Using Sensitivity Analyses in Knee Models. *The Journal of Arthroplasty*. 40. 10.1016/j.arth.2025.02.061.
- [2] Schlatterer B, Linares JM, Chabrand P, Sprael JM, Argenson JN. Influence of the optical system and anatomic points on computer-assisted total knee arthroplasty. *Orthop Traumatol Surg Res*. 2014 Jun;100(4):395-402. doi: 10.1016/j.otsr.2013.12.029. Epub 2014 May 14. PMID: 24835533.