



Multilevel Modeling of Resection Accuracy: Insights from 10,144 Clinical Cases using A Contemporary Computer-Assisted Total Knee Arthroplasty System

Yifei Dai^{1*}, Charlotte Bolch^{1*}, Amaury Jung², Cyril Hamad²

¹Exactech Inc, Gainesville, FL 32653, USA

²Blue Ortho, Gieres, FR

yifei.dai@exac.com; charlotte.bolch@exac.com

Abstract

This study applied an advanced statistical tool (multilevel modeling) to assess the accuracy of bony resection during total knee arthroplasty on 10144 cases performed using a modern CAOS system. An extensive list of factors was included for the modeling, including geographic region, inter-surgeon difference, surgeon's adoption of the technology (learning or proficient phases), and historical progression of the CAOS application (software versions). The comprehensive analysis demonstrated that the CAOS system is an accurate and precise solution to assist the surgeons to achieve his/her surgical resection goals.

1 Introduction

Accurate positioning of the implants is critical for the success of total knee arthroplasty (TKA) [1-4]. Numerous studies have confirmed the benefit of computer assisted orthopedic surgery (CAOS) in improving the accuracy of bony resection and limb alignment [5,6]. However, there are some common limitations shared across the existing studies that often fall into the following categories: 1) the studies were not powered enough to investigate geographic and inter-surgeon variance; 2) longitudinal performance of a specific CAOS system was overlooked, despite improvements and updates in the software applications are the standard practice in marketed systems.

It is unquestionably difficult to initiate clinical studies that encompass the clinical cases performed by a specific CAOS system with sufficient sample size for stratifying geographic regions, variation of usage between individual surgeons, and software updates over the system's application history. Nowadays, modern cloud-based data infrastructure allows archiving of technical data without the need

to assess patient information, providing possibilities to comprehensively assess the accuracy of a CAOS system across its users, geographic regions, and history of its application. This study aimed to apply an advanced statistical analysis (multilevel modeling) to assess resection accuracy across the entire TKA application history of a modern CAOS system. Specifically, the authors sought to determine the impact on accuracy from 1) geographic region; 2) inter-surgeon difference; 3) surgeon's adoption of the technology (learning curve); and 4) historical progression of the CAOS application (software versions).

2 Materials and Methods

A retrospective review and analysis of a proprietary cloud-based web that archives all TKAs performed using a modern CAOS system from its first application to the time of this study. All logs contained technical information recorded on the surgical cases. The database did not include patient information of any sort. Similarly, all surgeons were de-identified with only their geographic information available.

Resection errors (accuracy) in the proximal tibia and distal femur were evaluated in this study. Multilevel modeling was used to understand whether and where the variability is located in the resection errors in both tibia and femur. The errors were from two sources: 1) a level-1 variability that reflects resection errors across surgeries; and 2) a level-2 variability that describes the attribution of resection errors across a grouping variable, including geographic region, inter-surgeon differences across established surgeons, adoption phases (learning/proficient), and version of the CAOS software application (Table 1).

A total of 24 unconditional multilevel models were run to determine whether there were differences within each of the 4 level-2 grouping variables across six accuracy measurements (tibia: varus/valgus alignment, posterior slope, resection depth; and femur: varus/valgus alignment, flexion/extension, and resection depth). The model was formulated as following:

$$(1) \text{ Accuracy measurement}_{ijk} = \gamma_{00k} + U_{0jk} + \varepsilon_{ijk}$$

Where γ_{00k} = the grand mean of the resection error, U_{0jk} = random residual for level II variance, ε_{ijk} = random residual for level I variance, $U_{0j} \sim N(0, \sigma_{\mu_0}^2)$, $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon}^2)$, and i = individual cases, j = level-2 variable (ID of grouping categories), k = accuracy measurement for tibia (1) or femur (2). For each model, level-1 and level-2 variance estimates were used to compute an intraclass correlation (ICC). The ICC quantifies the proportion of variance at level-2 ranging between 0 (no variance) and 1 (100% of the variance). All models estimated used maximum likelihood estimation methods and were performed in SAS 9.3 (SAS Institute, Cary, NC, USA).

3 Results

A total of 10,144 CAOS TKA cases from December 2010 to August 2018 were reviewed. For the tibial resection, the deviations in coronal alignment (tibial varus/valgus angle) and sagittal alignment (posterior tibial slope) were $0.06 \pm 0.94^\circ$ and $-0.09 \pm 1.73^\circ$, respectively. For the femoral resection, the deviations in coronal alignment (femoral varus/valgus angle) and sagittal alignment (femoral flexion) were $0.00 \pm 0.97^\circ$ and $-0.17 \pm 1.44^\circ$, respectively.

ICC values are summarized in Table 2. Variation in geographic region, CAOS software application versions, and adoption phases (learning/proficient) all exhibited to account for negligible amounts of total variability in tibial and femoral resection errors (< 0.02). Notably, inter-surgeon differences accounted for between 0.0223 and 0.2444 of the total variability in tibia and femur resection errors, which was within the commonly acceptable natural variations in observational studies [7]. A further investigation of the inter-surgeon differences revealed that for the tibia, 100%, 97.6%, and 95.2% of the surgeons had less than $2^\circ/\text{mm}$ standard deviations in the resection errors of varus/valgus alignment,

posterior slope, and resection depth, respectively. Similarly for the femur, the percentages were 100%, 97.6%, and 97.6% for varus/valgus alignment, flexion/extension, and resection depth, respectively.

4 Discussion

Malpositioning of the implants negatively impacts the outcomes of TKA [1-4]. The study demonstrated high accuracy in bony resections by using the CAOS system. Furthermore, the resection accuracy was not sensitive to geographic region, CAOS software application version, or learning period. Although some variations were shown in established surgeons, the ICC values reported were within the established definition of commonly accepted variabilities from observational studies (ICC between 0.15 and 0.25 [7]). As such, no meaningful variability was observed from this study with regard to established surgeons.

To date, this is the first big data analysis applying advanced statistical modeling to assess the accuracy of a CAOS system across all its application history, extensively considering factors that may influence the bony resections. All, not just selective, surgeons, geographic regions, software versions, and phases of adoption were assessed, making this analysis an objective and comprehensive review of the accuracy performance of the system.

It has been questioned by many investigators that the accuracy in the alignment measured based on standard long-leg standing load-bearing radiograph may be compromised by the quality of the image, inter- and intra- observer variability, and can be sensitive to the position of the limb or direction of the beam that lead to an oblique (not strict anteroposterior) view. The intraoperative resection alignment check by directly pressing the instrumented checker on the bony resection surface provided a robust and consistent measurement of the bony resection alignment directly against the intraoperatively established alignment reference system.

Table 1. Grouping variables for the assessment of level-2 variability.

Grouping Variables	Definition of Categories	Number of Categories
Geographic Regions	APAC: Japan, Australia, Korea, Singapore, India EU: France, Switzerland, United Kingdom, Italy, Spain USA	3
Individual Established Surgeons	A data subset containing surgeons with $\geq 50^*$ cases experience with the CAOS system. Each individual surgeon was treated as one category.	41
Adoption Phases	Learning: combined cases #1-15 from all established surgeons Proficient: combined cases #36-50 from all established surgeons	2
Software Application	Software versions	6

* The selection of ≥ 50 cases to define established surgeons was based on consideration of maintaining sufficient sample size per category.

Table 2. ICC values for level-2 variables from multilevel models.

ICC Accuracy Measurement	Grouping Variables			
	Geographic Regions	Individual Established Surgeons	Adoption Phases	Software Versions
Tibial Resection Errors				
Var/Val Alignment	0.0000	0.0736	0.0000	0.0038
Posterior Slope	0.0037	0.1909	0.0000	0.0093
Resection Depth	0.0014	0.2444	0.0000	0.0073
Femoral Resection Errors				
Var/Val Alignment	0.0016	0.0223	0.0001	0.0003
Flexion/Extension	0.0069	0.1024	0.0197	0.0075
Resection Depth	0.0005	0.0707	0.0046	0.0019

5 References

- [1] Choong PF, et al. Does accurate anatomical alignment result in better function and quality of life? Comparing conventional and computer-assisted total knee arthroplasty. *J Arthroplasty* 24(4):560–9,2009.
- [2] Blakeney WG, et al. Functional outcomes following total knee arthroplasty: a randomised trial comparing computer-assisted surgery with conventional techniques. *Knee* 21(2):364–8,2014.
- [3] Huang NF, et al. Coronal alignment correlates with outcome after total knee arthroplasty: five-year follow-up of a randomized controlled trial. *J Arthroplasty* 27(9):1737–41,2012.
- [4] Longstaff LM, et al. Good alignment after total knee arthroplasty leads to faster rehabilitation and better function. *J Arthroplasty* 24(4):570–8,2019.
- [5] Brin YS, et al. Imageless computer assisted versus conventional total knee replacement. A Bayesian meta-analysis of 23 comparative studies. *Int Orthop*,35(3):331-9,2011.
- [6] Hetaimish BM, et al. Meta-analysis of navigation vs conventional total knee arthroplasty. *J Arthroplasty*,27(6):1177-82,2012.
- [7] Hedges, LV and Hedberg, EC. Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29(1), 60–87,2007.