Clinical Evaluation of Artificial Intelligence based Preoperative Plans for Total Knee Arthroplasty

Adriaan Lambrechts [1*†], Muthu Ganapathi[2*] and Roel Wirix-Speetjens[1]

[1] Materialise NV
[2] Ysbyty Gwynedd, Bangor, UK
Adriaan.lambrechts@materialise.be

Abstract

A preoperative plan is a virtual plan that defines the implant position and orientation allowing a surgeon to prepare for surgery. However, the default preoperative plan for total knee arthroplasty proposed by the manufacturer requires changes to be made by the surgeon in more than 90% of the cases. Previous studies have shown that artificial intelligence can be used to create better preoperative plans compared to manufacturer's default plans. However, the quality of artificial intelligence based preoperative plans has not yet been compared to surgeon approved preoperative plans. The purpose of this study is to compare default, artificial intelligence and surgeon approved preoperative plans, by having them scored on a range from 1 (totally unacceptable plan) to 5 (no corrections needed) by an experienced surgeon, while being blinded to the plan type. Through a Wilcoxon signed rank test with α=0.05, AI based preoperative plans were found to be a significant improvement upon the default plans (p-val=0.000136), while the differences in score between AI and surgeon approved preoperative plans were insignificant (p-val= 0.083). Consequently these results indicate that AI generated preoperative plans for TKA are an improvement upon current default plans, which could increase the surgeon's planning efficiency when applied in clinical practice.

1 Introduction

Patient specific instrumentation (PSI), navigation and robotic surgery are computer assisted surgical (CAS) techniques to obtain good knee alignment in total knee arthroplasty (TKA) [1]. If combined with 3D imaging, like MRI or CT, surgeons can plan the procedure prior to the surgery. This allows the surgeon to get an idea which implant sizes would be needed, how the implants should be rotated, and

---

[*] Shared co-first authorship
[†] Created the first draft of this document

how much bone should be resected. A preoperative plan, thus, entails 14 degrees-of-freedom: 6 implant rotation parameters, 6 implant translation parameters and 2 implant sizes. Currently surgeons obtain default preoperative plans for each case from the manufacturer of the CAS technology.

Previous studies have shown that these default preoperative plans are not always appropriate for use during surgery, and thus require corrections by the surgeon [2]. A study by Okada et al. found that 91% of preoperative plans required corrections [3]. These corrections are needed to avoid intra-operative deviations from the preoperative plan, which are difficult and time consuming. Surgeon corrected preoperative plans have been shown to accurately predict the implant component size needed during surgery [2].

In our previous study, an Artificial Intelligence (AI) algorithm was proposed to create patient- and surgeon-specific preoperative plans [4]. The AI based preoperative plans significantly improve upon default plans by reducing the number of changes required by the surgeon, by half. However, as the number of changes does not capture planning parameter coherency, a blinded, randomized case comparison was done between default plans and AI plans by scoring them on a scale of 1 to 5 [5]. A score of one indicates a preoperative plan which is totally unacceptable, while a score of five indicates a plan which needs no corrections by the surgeon. The default plans scored 1.8, while the AI plans scored 3.6. The studies had two limitations: (1) It only involved one surgeon, and (2) it did not compare the AI plans against the surgeon corrected plans, which are actually used during surgery. If AI plans score as good as surgeon corrected plans, these would be a better alternative to be used in clinical practice, than the currently proposed default plans.

Therefor the goal of this study is twofold, on the one hand we want to validate if the AI algorithm translates to a different surgeon and implant system, and on the other hand if AI plans score comparable with surgeon corrected plans.

2 Methods

To test our hypotheses, an AI planning algorithm was created based on 110 retrospectively collected TKA cases planned by the surgeon (MG). All patients were implanted with Zimmer-Biomet NexGen® implants. 90 Cases were used to train the AI model, while the remaining 20 were used for clinical validation. Based on the default plan, commonly used bony landmarks and measurements were extracted, on the basis of which the AI model predicts the surgeon corrected preoperative plan parameters. This model has two sequential steps: (1) a feature selection step using the LASSO algorithm, and (2) a set of support vector regression models that predict the planning parameters [6], [7].

Using the trained models, the AI preoperative plans were generated for the 20 test cases. Thus, for each of the 20 test cases, three preoperative plans were available: the default, AI, and surgeon-corrected plan (collected prior to the study). For each of the test cases the surgeon was simultaneously shown the three preoperative plans in randomized order while being blinded to the source algorithm. For each plan the surgeon had to score its quality on a scale of 1 to 5 [5]. A score of one indicates a preoperative plan which is considered unacceptable by the surgeon, while a score of five indicates a plan which needs no corrections.

To determine if the AI generated preoperative plans are an improvement over surgeon-corrected or default plans a Wilcoxon signed rank test was used with $\alpha = 0.05$. The sample size, namely 20 cases was determined in order to obtain $\beta = 0.20$, corresponding with a statistical power of 80%.

3   Results & discussion

The scores given for all preoperative plans can be found in Table 1. The default plan obtained a score of 2.9 ± 0.64 (mean $\pm$ standard deviation), while the AI plan was scored 4.5 ± 0.61 (Figure 1). In 18 cases the AI plan achieved a higher score than default plans, whereas in the remaining two cases both scored equally well. From the statistical testing we can conclude than AI plans are a significant improvement (p-val = 0.000136) over the default plans. As in a previous study we demonstrated that AI based preoperative plans are more surgeon- and patient-specific resulting in higher quality preoperative plans [5]. However this provides further proof that AI plans also generalize towards other surgeons and different implant systems.

The surgeon corrected plans scored 4.8 ± 0.41, slightly higher than the AI plans. However, the differences between AI and surgeon corrected plans were found to be statistically insignificant (p-val = 0.0832). In 8 cases the AI and surgeon corrected plan scored equally high, while in three cases the AI plan actually received a higher score. Hence, these promising results answer our second research questions: AI plans seem to be comparable in quality to surgeon corrected preoperative plans.

The correction time for a preoperative plan by a surgeon for a default plan is currently 8 ± 4 minutes [8]. Our results could indicate that AI based preoperative plans could reduce preoperative planning time, compared with current plans. However, further research is necessary to verify this hypothesis.
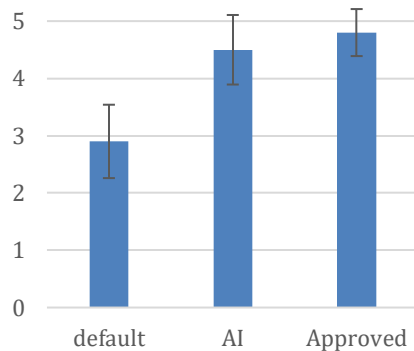

Figure 1: The average scores per preoperative planning type. The error bars indicate the standard deviations

Case Number	Default plan score	AI plan score	Surgeon corrected plan score
1	3	4	5
2	2	5	5
3	2	4	5
4	3	5	5
5	3	5	4
6	2	4	5
7	3	5	5
8	4	5	5
9	2	5	5
10	3	5	5
11	4	4	5
12	3	5	5
13	3	5	4
14	3	4	5

15	4	5	4
16	3	4	5
17	3	4	5
18	3	3	4
19	3	5	5
20	2	4	5

Table 1: The scores given by the surgeon per case and for each of the three preoperative plan types.

4 Conclusion

The present study had two objectives: firstly, analyzing if artificial intelligence based preoperative planning algorithms generalize towards different surgeons and implant systems, and secondly, to determine if AI plans are scored similarly as surgeon approved plans in terms of quality. To this aim we conducted a blinded randomized case comparison study investigating if surgeons score manufacturer's default plans, AI plans, and surgeon-corrected preoperative plans differently. As for our first research question, the AI based planning algorithm works for different surgeons with different implant systems. More promising, our results indicate that AI and surgeon corrected preoperative plans scored similarly, as determined by a Wilcoxon signed rank test. These findings might implicate that AI generated preoperative plans could reduce preoperative planning time associated with PSI based TKA surgery.

References

[1] K. Deep, S. Shankar, and A. Mahendra, "Computer assisted navigation in total knee and hip arthroplasty," *SICOT J*, vol. 3, p. 50, 2017, doi: 10.1051/sicotj/2017034.

[2] M. G. M. Schotanus, D. A. L. Schoenmakers, R. Sollie, and N. P. Kort, "Patient-specific instruments for total knee arthroplasty can accurately predict the component size as used peroperative," *Knee Surg Sports Traumatol Arthrosc*, vol. 25, no. 12, pp. 3844–3848, Dec. 2017, doi: 10.1007/s00167-016-4345-1.

[3] Y. Okada, A. Teramoto, T. Suzuki, Y. Kii, K. Watanabe, and T. Yamashita, "Preoperative corrections are required for planning of patient-specific instrumentation in total knee arthroplasty," *Knee*, vol. 24, no. 6, pp. 1492–1497, Dec. 2017, doi: 10.1016/j.knee.2017.09.008.

[4] A. Lambrechts, R. De Vloo, R. Wirix-Speetjens, and S. Van Huffel, "Using Artificial Intelligence to Improve Patient Specific Preoperative Plans for Total Knee Arthroplasty." Annual Meeting of the American Academy of Orthopaedic Surgeons (AAOS) proceedings, Mar-2019.

[5] A. Lambrechts, R. De Vloo, R. Wirix-Speetjens, and S. Van Huffel, "Artificial Intelligence based preoperative planning for TKA significantly improves current planning algorithms: a surgeon validation study." Annual Meeting of the American Academy of Orthopaedic Surgeons (AAOS) proceedings, Mar-2020.

[6] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[7] V. Vapnik, S. E. Golowich, and A. Smola, "Support Vector Method for Function Approximation, Regression Estimation and Signal Processing," in *Proceedings of the 9th International Conference on Neural Information Processing Systems*, Cambridge, MA, USA, 1996, pp. 281–287.

[8] M. Pietsch, O. Djahani, M. Hochegger, F. Plattner, and S. Hofmann, "Patient-specific total knee arthroplasty: the importance of planning by the surgeon," *Knee Surg Sports Traumatol Arthrosc*, vol. 21, no. 10, pp. 2220–2226, Oct. 2013, doi: 10.1007/s00167-013-2624-7.