Machine Learning with Arthroplasty-Specific Data: Results from Two German Registries at a Single Center

Igor Lazic[1], Florian Hinterwimmer[2], Ruediger Von Eisenhart-Rothe[3]
[1]Department of Orthopaedics and Sports Orthopaedics, Klinikum Rechts der Isar, [2]Institute for AI and Informatics in Medicine, Technical University of Munich, [3]Department of Orthopaedics and Sports Orthopaedics, Department For Orthopedics, Klinikum Rechts Der Isar, Technische Universität

INTRODUCTION: The identification of risk factors in primary and revision arthroplasties of the knee and hip currently relies on literature reviews and applying conventional statistical methods. Since the outcome prediction with risk stratification tools has been moderately accurate so far, a more comprehensive approach for clinical practice is needed. Machine Learning algorithms (ML) have had ample success in many disciplines. However, in orthopaedic research these methods have not yet had significant impact. In this study, we evaluate a standardized approach for ML in arthroplasty to predict complications and irregular surgery durations using data from two German arthroplasty-specific registries and investigate the potential of ML compared to conventional statistical analysis.

METHODS:

The dataset is based on two initiatives of the German Society for Orthopaedics and Orthopaedic Surgery (Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC)) to improve the quality of care. The German Arthroplasty Registry (Endoprothesenregister Deutschland (EPRD)) reports procedure and implant-related data of hip and knee replacements. EndoCert is a certification process of medical facilities in the field of joint replacement and is used to monitor compliance with structural, process and outcome quality standards in hospitals. We defined the problem statement and initial parameters. After screening, cleaning and preparation of these datasets, we gathered all cases from a single academic center with 12 different parameters from 2016 to 2019. 864 cases for primary and 596 cases for revision total knee arthroplasty (TKA), as well as 1217 cases for primary and 539 cases for revision total hip arthroplasty (THA) were included, respectively. We chose the XGBoost algorithm and applied a hyperparameter search, a cross validation and a loss weighting to cope with imbalanced and limited data. For final evaluation we calculated several metrics (accuracy, sensitivity, specificity, AUC).

RESULTS:

The metric results of XGBoost for prediction of complications and irregular surgery durations are summarized in Table 1. No relevant correlations could be shown using classical statistical methods, whereas the feature importance indicated multiple non-linear correlation. Age, height, weight, surgeon and diagnosis were found to be highly relevant parameters for prediction.

Table 1: Metric results of XGBoost for prediction of Complications and irregular surgery durations.

Complications	**Sensitivity**	**Specificity**	**Accuracy**	**ROC AUC**
Primary TKA	0.35	0.96	0.92	0.78
Revision TKA	0.59	0.94	0.91	0.82
Primary THA	0.31	0.84	0.64	0.8
Revision TKA	0.54	0.94	0.9	0.81
Irregular Surgery Duration	**Sensitivity**	**Specificity**	**Accuracy**	**ROC AUC**
Primary TKA	0.74	0.96	0.94	0.92
Revision TKA	0.9	0.85	0.87	0.95
Primary THA	0.58	0.92	0.82	0.89
Revision TKA	0.86	0.86	0.85	0.94

DISCUSSION AND CONCLUSION:

We identified risk factors for complications and irregular surgery durations after primary and revision arthroplasty of the hip and knee with a ML-model, that were not identified using conventional statistical analysis. An interdisciplinary interpretation as well as evaluation of the results by a data scientist and an orthopaedic surgeon are of paramount importance.

Complications	Sensitivity	Specificity	Accuracy	ROC AUC
Primary TKA	0.35	0.96	0.92	0.78
Revision TKA	0.59	0.94	0.91	0.82
Primary THA	0.31	0.84	0.64	0.8
Revision TKA	0.54	0.94	0.9	0.81
Irregular Surgery Duration	**Sensitivity**	**Specificity**	**Accuracy**	**ROC AUC**
Primary TKA	0.74	0.96	0.94	0.92
Revision TKA	0.9	0.85	0.87	0.95
Primary THA	0.58	0.92	0.82	0.89
Revision TKA	0.86	0.86	0.85	0.94

Table 1: Metric Results of XGBoost for Prediction of Complications and Irregular Surgery Durations.