A Deep Learning Algorithm for Predicting Postoperative Pain following Reverse Shoulder Arthroplasty Based on Preoperative Factors

Tim Schneller[1], Andrea Cina[4,5], Philipp Moroder[1], Markus Scheibel*[1,3], Asimina Lazaridou*[1,2]

[1]*Department for Shoulder and Elbow Surgery, Schulthess Clinic, Zürich, Switzerland*

[2]*Department of Anesthesiology, Brigham & Women's Hospital and Harvard Medical School, Boston, MA*

[3]*Center for Musculoskeletal Surgery, Charité-Universitätsmedizin, Berlin, Germany.*

[4]*Spine Center, Schulthess Clinic, Zürich, Switzerland*

[5]*Department of Health Sciences and Technology, ETH Zürich, Switzerland*

Abstract

Introduction

In shoulder arthroplasty, most research has predominantly concentrated on identifying optimal treatment strategies to enhance shoulder function, with comparatively less emphasis on postsurgical pain. Nevertheless, pain is an equally significant outcome in orthopedic surgery, especially in the geriatric population. The aim of this study was to develop a deep learning algorithm for prediction of postsurgical pain after reverse shoulder arthroplasty.

Methods

Clinical data from 1707 primarily treated patients from the local shoulder arthroplasty registry at Schulthess Clinic was extracted and used to build an artificial neural network. The model was set up with an input of 34 preoperative features. The target variable, a binary pain assessment derived from a numeric rating scale [0,10], was categorized as follows: if the pain scored 3 or higher, the record was labeled as belonging to the positive class; conversely, if the pain scored 2 or lower, the record was categorized as part of the negative class. The model was internally validated with a validation dataset that was comprised of 20% of the whole dataset. Model performance was evaluated on the validation dataset using the metrics accuracy, precision, recall, and f1-score.

Results

Overall, our model achieved an accuracy of 64% in predicting postsurgical pain two years after rTSA. Specificity was 56%/70%, while sensitivity was 60%/66% for the positive and negative class. The f1-score, a harmonic mean between specificity and sensitivity, was 58% and 68% for the positive and the negative class, respectively.

Discussion

The goal of developing an artificial neural network to predict postsurgical pain after reverse total shoulder arthroplasty with acceptable performance was reached successfully. However, more efforts are required to fine tune the model and further improve performance. In fact, other machine learning algorithms like random forest or boosted decision trees need to be considered as well. In a clinical setting, the availability of a prediction model tailored to the local patient population could help health providers and patients to make more informed decisions when it comes to deciding whether to perform reverse shoulder arthroplasty.

Introduction

Shoulder arthroplasty, a surgical procedure aimed at restoring function and alleviating pain in patients with various shoulder pathologies, has garnered considerable attention in last years. While numerous studies have rightfully emphasized functional outcomes following shoulder arthroplasty, the aspect of persistent postoperative pain has often been overlooked [1-4].

Despite advancements in surgical techniques and implant designs, a subset of patients continues to experience persistent pain even after successful shoulder arthroplasty. This persistent pain can significantly impact patients' quality of life, functional outcomes, and overall satisfaction with the procedure.

Irrespective of pain being an equally as important outcome in orthopedic surgery, prior work has focused primarily on functional outcomes such as range of motion, strength and patient-reported outcome measures, often neglecting the nuanced assessment of postoperative pain [5-7]. Consequently, the mechanisms underlying persistent pain following shoulder arthroplasty remain inadequately understood, hindering the development of targeted interventions to mitigate this issue.

Recognizing the importance of addressing persistent postoperative pain in patients undergoing shoulder arthroplasty is paramount for delivering comprehensive care. By expanding our attention beyond functional outcomes and integrating a focus on postoperative pain management, clinicians and researchers can more effectively meet the holistic needs of patients.

Leveraging advancements in machine learning, for instance, offers a promising avenue to predict postsurgical pain based on preoperative factors. In fact, employing machine learning techniques has become state of the art in medical research in recent years. Machine learning applications are manifold and pose an enormous opportunity for any medical field including orthopedic surgery, which should be acted upon [8-11].

By identifying high-risk individuals and tailoring interventions accordingly, we can enhance patients' surgical experiences and long-term outcomes, ultimately advancing the field of shoulder arthroplasty and improving patient care. Past studies by Kumar et al utilizing several different machine learning algorithms demonstrated impressive accuracy rates, ranging from 93% to 99% for patient-reported outcome measures (PROMs) and 85% to 94% for function and range of motion (ROM) measures, in accurately identifying patients who achieved clinical improvement surpassing the minimal clinically important difference (MCID) [12-14]. Other work by Franceschetti et al. showed that machine learning can predict anterior elevation after rTSA with an accuracy of 88% using a Support Vector Machine [15]. Similarly, Oeding et al. developed a model with the capability to predict readmission within 90 days following total shoulder arthroplasty using Extreme Gradient Boosting with a AUROC of 0.7 [16].

Therefore, there is a call for identifying preoperative factors that may serve as reliable indicators of postoperative pain intensity and duration. Understanding these factors can aid in personalized pain management strategies and improve patient care and clinical practice.

The main objective of this study was to develop a deep learning algorithm for predicting postsurgical pain after reverse shoulder arthroplasty (rTSA). A secondary objective is the identification of possible indicators of pain following rTSA.

Methods

Clinical and self-reported data from patients treated with a reverse total shoulder arthroplasty between March 2006 and February 2022 from our local shoulder arthroplasty registry were extracted. This data served as input for a deep artificial neural network (ANN) used to predict pain levels two years post-surgery. The study was performed in accordance with the standards of the Ethics Committee of Zurich (KEK-ZH-Nr. 2014-0483) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All patients provided written informed consent.

Outcome of Interest

The outcome of interest in this analysis was postoperative pain defined as persistent pain after shoulder arthroplasty 2 years postoperatively.

The outcome variable, calculated from the pain Numeric Rating Scale (NRS) (0 = no pain, 10 = worst pain) was categorized as follows: if the pain scored 3 or higher, the record was labeled as belonging to the positive class; conversely, if the pain scored 2 or lower, the record was categorized as part of the negative class. Pain was assessed preoperatively and postoperatively at the 2 years follow up using a self-report patient questionnaire.

Features

The following features were included in the analysis: patient-related and sociodemographic covariates including patient age at the time of surgery, sex, baseline smoking status (yes, no), baseline alcohol consumption status (no, occasional, daily), Insurance type (general, semi-private, private) and the American Society of Anesthesiologists (ASA) physical status classification system [17]. To detect any potential non-linear relationships, age squared was also included as a feature.

Disease-related covariates included dominance of the affected shoulder, bilateral shoulder arthroplasty (yes or no), surgery admission type (illness or accident), primary surgery diagnosis (Cuff Tear Arthropathy (CTA), rheumatoid arthritis, fracture sequelae, primary osteoarthritis (OA), primary humeral head necrosis), previous surgeries (yes or no), comorbidities (yes or no) and performed biceps tenotomy o during surgery (yes or no).

Objective data such as abduction strength and which type of implant was implanted (Promos Reverse, Lima Reverse, Univers Revers, Aequalis Reverse II, Ascend Flex Reverse or Perform humeral reverse) was used as input to the model. Additionally, we included three indexes of active range of motion (AROM): flexion, abduction and external rotation with arms comfortably at the side. As for PROMS, we included baseline Shoulder Pain and Disability Index (SPADI) divided into function and pain subscores [18], baseline pain assessed on a numeric rating scale and the baseline short version of the Disability of the Arm, Shoulder and Hand Questionnaire (quickDASH) [19].

Eligibility criteria

Inclusion criteria included: patients that underwent reverse shoulder arthroplasty between 2006-2022, patients who provided consent for use of their data for research purposes.

Exclusion criteria included: patients with acute fracture and revision indications were excluded from the dataset.

After performing all the above steps, a total of 1707 patients were included in our model, which was set up based on a selection of preoperative features (Table 1). The input data was preprocessed as follows: Nominal input features were dichotomized using one-hot-encoding. Continuous and ordinal variables were transformed by subtracting the mean and scaling to unit variance. Scaling features is a common practice in machine learning preprocessing to avoid variables of higher scale (e.g. age) being weighted more strongly, hence biasing the model prediction. A common challenge in scaling is data leakage from the validation to the training set. To prevent this, the dataset was divided into training and validation set before fitting the Scaler object to the training set. Using the fitting parameters from the training set, the validation set was then transformed accordingly. This way, no information from the validation set could leak into the training set.

The model was internally validated with a validation dataset that was comprised of 20% of the whole dataset. The model's performance was evaluated on the validation dataset using the metrics accuracy, precision/specificity, recall/sensitivity, and f1-score. The artificial neural network was created using the software Python (version 3.11.4) [20] with the PyTorch

Framework (version 2.1.2) [21] and the libraries NumPy (version 1.24.3) [22], pandas (version 1.5.3) [23], Scikit-learn (version 1.2.2) [24] and matplotlib (version 3.7.1) [25].

The proposed artificial neural network was built with an input layer of 30 neurons, followed by three fully connected hidden layers of 128, 128 and 64 neurons each. Each neuron of each layer was followed by a Rectified Linear Unit (ReLU) activation function, only the last layers neurons' output was passed through a sigmoid function to end up with the probability of belonging to the positive class. In total, the proposed artificial neural network had approximately 29313 trainable parameters.

When splitting the dataset into training and validation sets the stratified partitioning method was applied. The advantage of using this method is that both the training and validation sets each have the same class distribution, hence reducing bias and improving model performance.

Training of the model was performed using several mini batches of 32 observations per batch in parallel, which increases computational efficiency.

Minimization of the loss function (Binary Cross Entropy Loss) was done iteratively by calculating and updating weights and biases after passing each batch through the model using stochastic gradient descent and back propagation, performed by the "Adam" optimizer (learning rate=0.0001).

The validation set loss function started converging at 15 epochs, after which the training of the model was stopped. As soon as the loss function starts converging, the model does not learn much anymore, however, it starts overfitting, resulting in a very poor generalization of the model to unseen data.

A measure against class imbalance was taken by applying the SMOTE [26] algorithm, which stands for synthetic minority oversampling technique and augments the training set by creating linear interpolations of datapoints from underrepresented classes, essentially creating artificial data based on the training set. This results in a perfectly balanced training set with the same number of samples for each class, enabling the deep learning model to make more accurate predictions.

The analysis of features driving the prediction was performed using the SHAP (SHapley Additive exPlanations) library [27]. This method is based on game theory, where each feature of each observation is assigned a SHAP value, which is a calculated weighted contribution of each input feature to the prediction, allowing for easy interpretation of prediction-driving factors. However, these factors are not necessarily causative, they are merely the driving factors of the models' prediction, based on the available input data.

Results

In total, 1707 patients were used to train and internally validate an ANN which predicts postsurgical pain. The Dataset was split into training and validation sets, each consisting of 80% (1365 patients) and 20% (348 patients) of the entire dataset, respectively. SMOTE was employed on the training set, which resulted in balanced classes. After employing SMOTE, the augmented training set consisted of 1606 patients (each class being represented by 803 patients).

Our proposed ANN effectively learns from preoperatively available data. Overall, the model achieved a prediction accuracy of 63%, with slightly better performance in the negative class across all performance metrics (see Table 2). Precision ranged from 56 to 70%, Recall from 60 to 66%, and the F1-score ranged from 58 to 68%.

SHAP Values

SHAP analysis revealed the driving features of the models' predictions. Figure 1 shows the features which have the biggest impact on the model output. Blue dots represent low feature values, while red dots stand for high feature values. In dichotomized features, such as sex, the blue and red dots symbolize zero and one (female and male in the case of sex, but it could mean no and yes for one-hot encoded variables). For continuous variables like the quickDASH score the blue dots denote patients scoring lower on the scale and red dots portray patients on the higher end of the scale. The same applies for ordinal variables like shoulder flexion. Again, blue dots represent lower values, while red dots depict higher values of the respective feature.

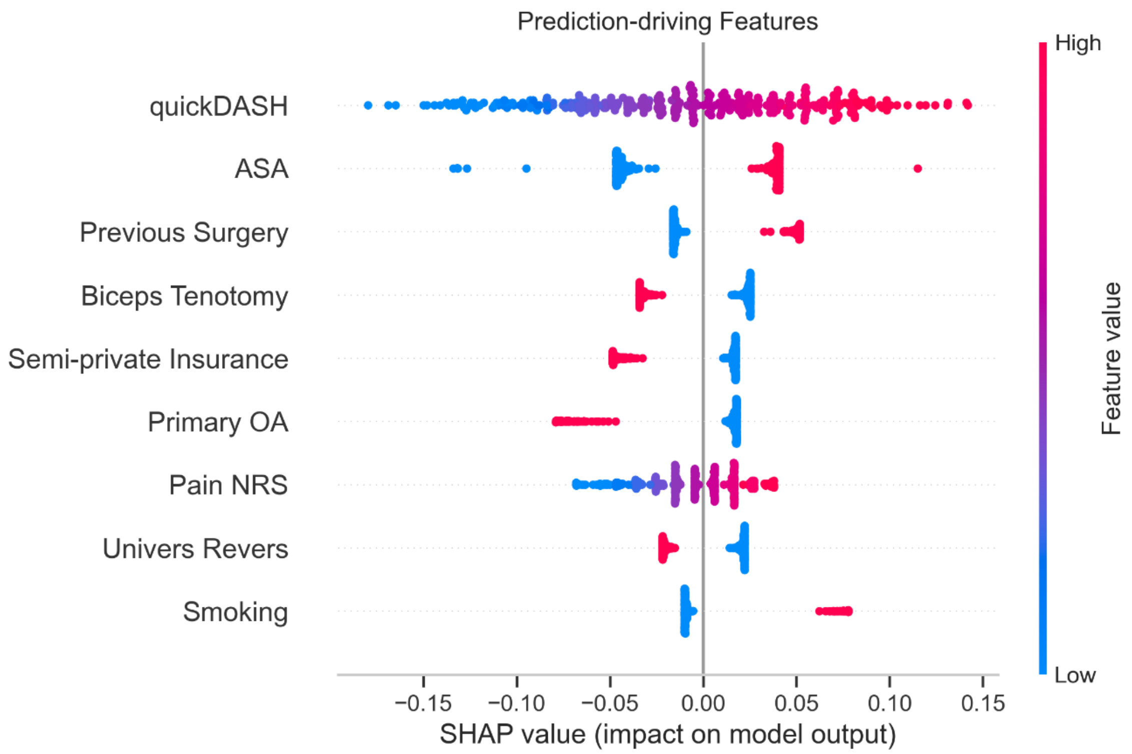

Figure 1: most impactful features on model output based on SHAP values.

The two most important features driving the prediction of postsurgical pain were the preoperative quickDASH score, followed by the ASA classification. Patients showing low preoperative quickDASH scores and low ASA grades tended to not develop postsurgical pain. Third, having had previous surgeries increased the likelihood of postsurgical pain, compared to those for whom the prosthesis was the first procedure on the affected joint. When the surgery included a biceps tenotomy, patients experienced lower postsurgical pain. Furthermore, having semi-private insurance status, being diagnosed with primary osteoarthritis, having a low preoperative pain level or receiving the Univers-Revers arthroplasty model are all indicators of not developing postsurgical pain. Smokers tended to higher postoperative pain levels. Higher baseline SPADI function subscores seemed to predict lower pain levels. Other protective factors of postsurgical pain included functioning external rotation, female sex, and lower age. Being admitted to surgery due to illness is indicative of lower pain levels post-surgery, compared to being admitted due to trauma. In comparison to other range of motion parameters, a high preoperative abduction indicates the occurrence of postsurgical pain.

An explanation of predictions on a case-level is depicted in Figure 2. It shows individual characteristics of selected patients, that are most important for the outcome prediction. For example, Patient B is predicted to experience no postsurgical pain despite not having had a tenotomy of the long biceps head. However, this patient also showed a low preoperative quickDASH score, a semi-private insurance status, a low disease burden as measured by the ASA grade, and low preoperative pain levels, none of them being indicative of developing postsurgical pain.

Figure 2: Individual predictions for selected patients.

Discussion

The main objective of this study was to investigate, whether leveraging preoperative data from a local shoulder arthroplasty registry would allow us to develop a predictive model using a deep learning algorithm for postsurgical pain following reverse shoulder arthroplasty. Our model demonstrated the ability to predict postsurgical pain solely based on preoperative factors, allowing for identification of preoperative risk-factors relating to postsurgical pain. Notably, compared to the baseline model, which is a uniform dummy classifier, our deep learning approach demonstrates superior performance, as shown in the confusion matrices for both models in Figure 3.

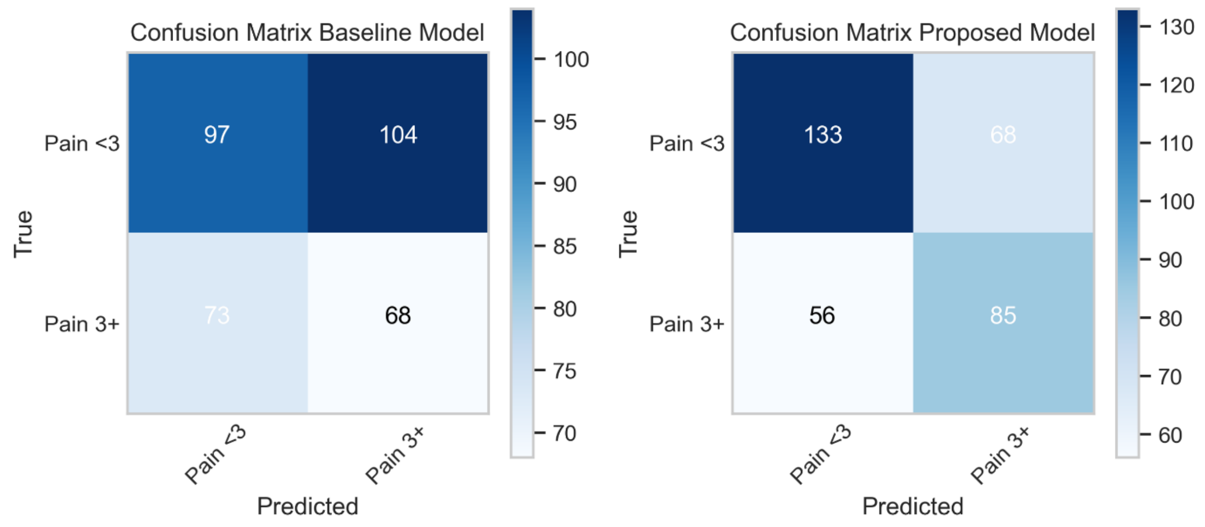

Figure 3: Confusion matrices for the baseline model and the proposed model.

The analysis of Shapley values allowed us to identify the features most relevant to the prediction of pain. The highest-ranking predictive feature for pain two years after rTSA was the quickDASH score, with lower baseline scores indicating lower postsurgical pain levels. These results come in agreement with a recent systematic review where patients with higher preoperative function are associated with better postoperative outcomes after shoulder arthroplasty [28]. While the quickDASH score neither is specific to the shoulder joint nor to the affected side, studies have shown that it correlates well with other scores such as the SF-36, SST, SPADI or ASES, demonstrating its validity in capturing the functional health status in patients with shoulder complaints [29, 30].

The second most predictive feature is the American Society of Anesthesiologists (ASA) physical status classification system. In our model, an ASA score of I or II seems to be associated with less postsurgical pain in contrast to ASA scores of III or IV. This finding comes in agreement with other studies demonstrating that better physical status preoperative predicts better outcomes in orthopedics surgeries. In fact, an analysis based on data of the New Zealand Joint registry showed that patients with ASA I and II have significantly increased postoperative Oxford scores, compared to ASA II or higher after hip arthroplasty. Furthermore, according to the authors findings, these patients have a lower mortality rate and are less likely to need an early revision [31].

Previous surgeries on the affected shoulder seemed to lead to more postsurgical pain, which agrees with prior evidence, where previous ipsilateral surgeries have been identified as a predictor of poor outcomes [32-34].

According to our models' predictions, a concomitant biceps tenotomy leads to less postsurgical pain. The role of long head of the biceps in causing shoulder pain has been a subject of controversy for many years. Extensive literature discusses its anatomy, function, pathology, and, most importantly, the appropriate treatment options [35]. Work by Boileau et al. demonstrated that isolated biceps tenotomy significantly alleviates pain in patients with massive irreparable rotator cuff tears [36]. Recent efforts by Veen et al. have confirmed these findings, showing a reduction in pain after an isolated tenotomy of the long head of the biceps in 78% of patients suffering from degenerative rotator cuff tears [37]. While these findings are rather conclusive for rotator cuff repair and seem to indicate a pain-relieving effect of the biceps tenotomy, the literature is scarce regarding the role of biceps tenotomy on postsurgical pain in shoulder arthroplasty. The results of our model are a first step towards filling this gap, as we demonstrate similar findings in reverse shoulder arthroplasty.

In the literature, a higher socioeconomic status is generally associated with and better clinical outcomes after joint replacement [38, 39]. For the shoulder joint, Sheth et al. analyzed the influence of socioeconomic factors on outcomes after anatomic shoulder arthroplasty and found that individuals with lower socioeconomic status (SES) were associated with increased postsurgical pain and lower function compared to individuals with higher SES [40]. For shoulder arthroplasty, studies report conflicting results on postoperative pain as there seems to be a difference between anatomic and reverse shoulder arthroplasty when it comes to postsurgical pain. Sheth et al. analyzed the influence of socioeconomic factors on outcomes after anatomic shoulder arthroplasty and found that individuals with lower socioeconomic status (SES) were associated with increased postsurgical pain and lower function compared to individuals with higher SES [40]. An article by Strotman et al. reported similar findings, showing that the outcome for Medicaid patients is inferior to patients with private insurance in shoulder arthroplasty [41]. In contrast, an article by Waldrop et al., although showing significant differences in pain between socioeconomic groups in the general shoulder arthroplasty cohort, failed to report any difference in pain between patients of higher and lower SES in the reverse shoulder arthroplasty cohort [42]. For our proposed model, contrary to the literature, private insurance status had a negligible effect on pain after reverse shoulder arthroplasty. In fact, patients with semi-private insurance tended to experience less postsurgical pain.

If the diagnosed pathology for shoulder arthroplasty was primary OA, patients developed less postsurgical pain according to our model. This finding is backed by the literature. In specific, Forlizzi et al. found that primary OA among many factors, was strongest associated with excellent clinical outcomes after rTSA, including pain [32]. A systematic review by Kennedy et al. showed that the reduction in pain after rTSA was greatest in patients diagnosed with primary OA [43].

Another feature, which according to our model leads to less postsurgical pain is the presence of the implant model Univers-Revers. While it has been shown that this specific implant delivers excellent clinical results including a significant reduction in pain [44], the finding should be interpreted with caution, as it is unprecedented in the literature to our knowledge. Further efforts are needed to verify this result, as there might be a selection bias in this cohort, since a lot of patients from this cohort were treated by a single surgeon that tended to treat more privately insured patients than others. As discussed above, the literature suggests that privately insured patients achieve better outcomes than patients with a lower SES, hence leading to better outcomes in patients treated with a Univers-Revers prosthesis.

A low preoperative pain level indicated low postsurgical pain according to our model, which reflects the literature. In fact, a systematic review by Hernandez et al., showed a strong association between postsurgical pain and high preoperative pain [45]. However, we cannot conclude from this finding that low preoperative pain is associated with low postsurgical pain. We can however say that according to this review low preoperative pain is not associated with postsurgical pain, which supports the models' prediction.

In our model, smokers tended to experience more postsurgical pain. This finding is in agreement with prior studies reporting worse outcomes after shoulder arthroplasty for smoking patients. Mechanistically, this is likely a causative factor, as wound and bone healing impairment due to tobacco is quite well-researched [46].

The main strength of this study is the large sample size including demographic, clinical, surgical and PROMS data, which to our knowledge has never been used to research postsurgical pain in the shoulder arthroplasty domain. A further strong point is leveraging this large dataset through the use of a sophisticated deep learning algorithm, potentially revealing findings that might not be apparent when using traditional statistical methods.

Providing surgeons with access to a tool based on our model could enhance decision-making processes, enabling more personalized interventions for patients. This approach could improve surgical outcomes and reduce healthcare costs, as patients at high risk of developing postsurgical pain could likely be treated with more alternative methods instead of reverse shoulder arthroplasty.

Limitations

Several limitations should be considered when interpreting results. First, the Numerical Rating Scale (NRS) used for pain assessment may not fully capture the complexity and variability of pain experiences. Second, several important predictors, such as medication use, body mass index (BMI), and the Subjective Shoulder Value (SSV), were not available, which could impact the robustness of the predictive model. Third, the lack of external validation might prevent the generalization of the model findings. Additionally, the presence of missing data, a common issue in registry-based studies, introduces potential bias that could affect the validity of the results. Finally, the decision not to explore machine learning algorithms other than deep neural networks may have resulted in the omission of potentially more effective algorithms such as Random Forest of Gradient Boosting.

Conclusion

This study demonstrates how machine learning modeling can offer valuable insights into the prediction of postsurgical pain following reverse total shoulder arthroplasty. By leveraging advanced algorithms, we were able to identify potential predictors and patterns that may not be immediately apparent through traditional statistical methods. However, while these findings are promising, they also highlight the need for further research. Continued efforts are essential to refine these models, validate their predictive accuracy across diverse patient populations, and incorporate additional relevant variables. Further machine learning algorithms should also be considered to optimize predictive ability. Such research will be crucial in ensuring that machine learning tools can be effectively integrated into clinical practice to improve patient outcomes.

References

1. Freislederer, F., et al., *Analysis of three different reverse shoulder arthroplasty designs for cuff tear arthropathy - the combination of lateralization and distalization provides best mobility.* BMC Musculoskelet Disord, 2024. **25**(1): p. 204.
2. Imiolczyk, J.P., et al., *Metallic humeral and glenoid lateralized implants in reverse shoulder arthroplasty for cuff tear arthropathy and primary osteoarthritis.* JSES Int, 2022. **6**(2): p. 221-228.
3. Imiolczyk, J.P., et al., *Lateralization and distalization shoulder angles do not predict outcomes in reverse shoulder arthroplasty for cuff tear arthropathy.* J Shoulder Elbow Surg, 2024. **33**(1): p. 121-129.
4. Moroder, P., et al., *Scapulothoracic orientation has a significant influence on the clinical outcome after reverse total shoulder arthroplasty.* J Shoulder Elbow Surg, 2024.
5. Endell, D., et al., *Impact of Sports Activity on Medium-Term Clinical and Radiological Outcome after Reverse Shoulder Arthroplasty in Cuff Deficient Arthropathy; An Institutional Register-Based Analysis.* J Clin Med, 2021. **10**(4).
6. Freislederer, F., et al., *Lateralized vs. classic Grammont-style reverse shoulder arthroplasty for cuff deficiency Hamada stage 1-3: does the design make a difference?* J Shoulder Elbow Surg, 2022. **31**(2): p. 341-351.
7. Grob, A., et al., *Glenoid Component Loosening in Anatomic Total Shoulder Arthroplasty: Association between Radiological Predictors and Clinical Parameters-An Observational Study.* J Clin Med, 2021. **10**(2).
8. Cabitza, F., A. Locoro, and G. Banfi, *Machine Learning in Orthopedics: A Literature Review.* Front Bioeng Biotechnol, 2018. **6**: p. 75.
9. Habehh, H. and S. Gohel, *Machine Learning in Healthcare.* Curr Genomics, 2021. **22**(4): p. 291-300.
10. Pruneski, J.A., et al., *Supervised machine learning and associated algorithms: applications in orthopedic surgery.* Knee Surg Sports Traumatol Arthrosc, 2023. **31**(4): p. 1196-1202.
11. Ogink, P.T., et al., *Wide range of applications for machine-learning prediction models in orthopedic surgical outcome: a systematic review.* Acta Orthop, 2021. **92**(5): p. 526-531.
12. Kumar, V., et al., *Using machine learning to predict clinical outcomes after shoulder arthroplasty with a minimal feature set.* J Shoulder Elbow Surg, 2021. **30**(5): p. e225-e236.

13. Kumar, V., et al., *What Is the Accuracy of Three Different Machine Learning Techniques to Predict Clinical Outcomes After Shoulder Arthroplasty?* Clin Orthop Relat Res, 2020. **478**(10): p. 2351-2363.
14. Roche, C., et al., *Validation of a machine learning-derived clinical metric to quantify outcomes after total shoulder arthroplasty.* J Shoulder Elbow Surg, 2021. **30**(10): p. 2211-2224.
15. Franceschetti, E., et al., *Machine learning can predict anterior elevation after reverse total shoulder arthroplasty: A new tool for daily outpatient clinic?* Musculoskelet Surg, 2024.
16. Oeding, J.F., et al., *Understanding risk for early dislocation resulting in reoperation within 90 days of reverse total shoulder arthroplasty: extreme rare event detection through cost-sensitive machine learning.* J Shoulder Elbow Surg, 2023. **32**(9): p. e437-e450.
17. Daabiss, M., *American Society of Anaesthesiologists physical status classification.* Indian J Anaesth, 2011. **55**(2): p. 111-5.
18. Breckenridge, J.D. and J.H. McAuley, *Shoulder Pain and Disability Index (SPADI).* J Physiother, 2011. **57**(3): p. 197.
19. Gummesson, C., M.M. Ward, and I. Atroshi, *The shortened disabilities of the arm, shoulder and hand questionnaire (Quick DASH): validity and reliability based on responses within the full-length DASH.* BMC Musculoskeletal Disorders, 2006. **7**(1): p. 44.
20. *Python Software Foundation. Python Language Reference, version 3.11.4, available at https://www.python.org/doc/versions/.*
21. Paszke, A., et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 2019.
22. Charles R. Harris, K.J.M., Stefan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernandez del Rio, Mark Wiebe, Pearu Peterson, Pierre Gerard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, Travis E. Oliphant, *Array programming with NumPy.* Nature, 2020. **585**: p. 357-362.
23. *pandas-dev/pandas: Pandas*. 2023, Zenodo. p. The pandas development team.
24. Pedregosa F., V.G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E., *Scikit-learn: Machine Learning in Python.* Journal of Machine Learning Research, 2011. **12**: p. 2825-2830.
25. Hunter, J.D., *Matplotlib: A 2D graphics environment.* Computing in Science & Engingeering, 2007. **9**: p. 90-95.
26. Chawla Nitesh, B.K., Hall Lawrence, Kegelmeyer Philip, *SMOTE: synthetic minority over-sampling technique.* Journal of artificial intelligence research, 2002. **16**: p. 321-57.
27. Scott M. Lundberg, S.-I.L., *A Unified Approach to Interpreting Model Predictions*, in *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA*. 2017.
28. Crutsen, J.R.W., et al., *Predictive factors for postoperative outcomes after reverse shoulder arthroplasty: a systematic review.* BMC Musculoskelet Disord, 2024. **25**(1): p. 439.
29. Roy, J.S., J.C. MacDermid, and L.J. Woodhouse, *Measuring shoulder function: a systematic review of four questionnaires.* Arthritis Rheum, 2009. **61**(5): p. 623-32.
30. SooHoo, N.F., et al., *Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36.* J Hand Surg Am, 2002. **27**(3): p. 537-41.
31. Hooper, G.J., et al., *The relationship between the American Society Of Anesthesiologists physical rating and outcome following total hip and knee arthroplasty: an analysis of theNew Zealand Joint Registry.* J Bone Joint Surg Am, 2012. **94**(12): p. 1065-70.

32. Forlizzi, J.M., et al., *Predictors of poor and excellent outcomes after reverse total shoulder arthroplasty.* J Shoulder Elbow Surg, 2022. **31**(2): p. 294-301.
33. Carducci, M.P., Z.R. Zimmer, and A. Jawa, *Predictors of unsatisfactory patient outcomes in primary reverse total shoulder arthroplasty.* J Shoulder Elbow Surg, 2019. **28**(11): p. 2113-2120.
34. Frank, R.M., et al., *Shoulder Arthroplasty Outcomes After Prior Non-Arthroplasty Shoulder Surgery.* JB JS Open Access, 2018. **3**(3): p. e0055.
35. AlQahtani, S.M. and R.T. Bicknell, *Outcomes following long head of biceps tendon tenodesis.* Curr Rev Musculoskelet Med, 2016. **9**(4): p. 378-387.
36. Boileau, P., et al., *Isolated arthroscopic biceps tenotomy or tenodesis improves symptoms in patients with massive irreparable rotator cuff tears.* J Bone Joint Surg Am, 2007. **89**(4): p. 747-57.
37. Veen, E.J.D., et al., *Arthroscopic isolated long head of biceps tenotomy in patients with degenerative rotator cuff tears: mid-term clinical results and prognostic factors.* Eur J Orthop Surg Traumatol, 2021. **31**(3): p. 441-448.
38. Clement, N.D., et al., *Socioeconomic status affects the early outcome of total hip replacement.* J Bone Joint Surg Br, 2011. **93**(4): p. 464-9.
39. Hoelen, T.A., et al., *Socioeconomic Status Affects Patient-Reported Outcome Measures in Total Hip and Knee Arthroplasty: A Retrospective Dutch Registry Study.* J Arthroplasty, 2024.
40. Sheth, M.M., et al., *Lower Socioeconomic Status Is Associated With Worse Preoperative Function, Pain, and Increased Opioid Use in Patients With Primary Glenohumeral Osteoarthritis.* J Am Acad Orthop Surg, 2020. **28**(7): p. 287-292.
41. Strotman, P., et al., *Effect of Insurance Status on Clinical Outcomes After Shoulder Arthroplasty.* Orthopedics, 2020. **43**(6): p. e523-e528.
42. Waldrop, L.D., 2nd, et al., *The effect of lower socioeconomic status insurance on outcomes after primary shoulder arthroplasty.* J Shoulder Elbow Surg, 2018. **27**(6S): p. S35-S42.
43. Kennedy, J., et al., *Reverse total shoulder arthroplasty clinical and patient-reported outcomes and complications stratified by preoperative diagnosis: a systematic review.* J Shoulder Elbow Surg, 2021. **30**(4): p. 929-941.
44. Schwyzer, H.K., et al., *Short-term safety, function, and quality of life in patients treated with Univers Revers prosthesis: a multicenter 2-year follow-up case series.* J Shoulder Elbow Surg, 2020. **29**(11): p. 2282-2291.
45. Hernandez, C., et al., *Pre-operative Predictive Factors of Post-operative Pain in Patients With Hip or Knee Arthroplasty: A Systematic Review.* Reumatol Clin, 2015. **11**(6): p. 361-80.
46. Hernigou, J. and F. Schuind, *Tobacco and bone fractures: A review of the facts and issues that every orthopaedic surgeon should know.* Bone Joint Res, 2019. **8**(6): p. 255-265.

Table 1

Parameter	**n**	**Mean (SD)**	**Median (range)**
Age	1707	73.8 (8.3)	74.6 (23.7;95.1)
Sex (n,%)			
Female	1140 (67)		
Male	567 (33)		
ASA Classification (n,%)			
Category 1	32 (2)		

Category 2	819 (48)		
Category 3	841 (49)		
Category 4	15 (1)		
Admission Type (n,%)			
Illness	1283 (75)		
Accident	424 (25)		
Insurance (n,%)			
General	663 (39)		
Semi-private	522 (31)		
Private	522 (31)		
Diagnosis (n,%)			
Cuff Tear Arthropathy (CTA)	1199 (70)		
Osteoarthritis due to Instability	2 (0)		
Rheumatoid Arthritis	59 (3)		
Fracture Sequelae	177 (10)		
Primary Osteoarthritis (OA)	247 (14)		
Primary Humeral Head Necrosis	23 (1)		
Previous Surgery (n,%)			
No	1155 (68)		
Yes	552 (32)		
Biceps Tenotomy			
No	1044 (61)		
Yes	663 (39)		
Implant model (n,%)			
Promos Reverse	512 (30)		
Lima Reverse	105 (6)		
Univers Revers	721 (42)		
Aequalis Reverse II	213 (12)		
Ascend Flex Reverse	118 (7)		
PERFORM Humeral Reverse	31 (2)		
Other	7 (0)		
quickDASH	1707	53.0 (18.2)	54.5 (4.5;97.7)
Shoulder Flexion	1707	79.0 (37.0)	75.0 (0.0;180.0)
Shoulder Abduction	1707	68.5 (31.6)	60.0 (0.0;180.0)
Shoulder External Rotation (0°)	1707	24.6 (19.7)	20.0 (-30.0;90.0)
Comorbidities (n,%)			
No	452 (26)		
Yes	1255 (74)		
Smoker (n,%)			
No	1542 (90)		
Yes	165 (10)		
Alcohol consumption (n,%)			
No	440 (26)		
Yes	881 (52)		
Occasionally	340 (20)		
Daily	46 (3)		
Age squared	1707	5521.7 (1168.2)	5565.2 (561.7;9044.0)
Pain NRS	1707	5.9 (2.5)	6.0 (0.0;10.0)
SPADI pain score	1707	34.2 (20.3)	31.8 (0.0;100.0)
SPADI function score	1707	37.6 (21.5)	34.1 (0.0;100.0)
Constant Score (power subscore)	1707	1.2 (2.9)	0.0 (0.0;22.0)

Pain postop above NRS 3 (n,%)			
No	1002 (59)		
Yes	705 (41)		

Table 2

Class	**Precision**	**Recall**	**F1-Score**	**N**
Pain below 3	0.70	0.66	0.68	201
Pain above or equal 3	0.56	0.60	0.58	141