Towards streamlining orthopedic consultations

Machine learning classification of knee diagnosis groups via computer-assisted history taking

Oosterhoff, Jacobien H.F.; Slaats, Twan; Warren, Tristan; van der Weegen, Walter

DOI
10.1016/j.knee.2025.05.017

Publication date
2025

Document Version
Final published version

Published in
Knee

The Knee 56 (2025) 241–248

Contents lists available at ScienceDirect

The Knee

journal homepage: www.elsevier.com/locate/thekne

Towards streamlining orthopedic consultations: Machine learning classification of knee diagnosis groups via computer-assisted history taking ☆

Jacobien H.F. Oosterhoff [a,*], Twan Slaats [a], Tristan Warren [b], Walter van der Weegen [c]

[a] Faculty of Technology Policy and Management, Delft University of Technology, Delft, the Netherlands
[b] InteractiveStudios, Den Bosch, the Netherlands
[c] Department of Orthopedic Surgery, Sports & Orthopedics Research Centre, St Anna Hospital, Geldrop, the Netherlands

ARTICLE INFO

Article history:
Received 16 January 2025
Revised 22 April 2025
Accepted 13 May 2025

Keywords:
Machine learning
Knee diagnosis
Diagnostic decision support
Computer-assisted history taking
Operational support

ABSTRACT

Background: The number of patients suffering from knee complaints is increasing, leading to increased orthopedic healthcare consumption. Predicting knee diagnoses prior to consultation may be valuable in optimizing the consultation workflow. Therefore, the purpose of this study was to develop and internally validate a machine learning (ML) algorithm for predicting a knee diagnosis group for patients aged 18 years and older, based on computer-assisted history taking.
Methods: A prospective cohort study at a single general district hospital was conducted to identify patients referred to an orthopedic surgeon for knee complaints. In total, 1172 patients were included, with an average age of 54 years (interquartile range 36–66), of which the majority were female (n = 594, 50.7%). The most frequent diagnosis group was knee osteoarthritis (n = 775, 66.1%), followed by ligamentous injuries (n = 208, 17.7%) and otherwise classified (n = 189, 16.1%). First, the dataset was randomly split 80:20 into training and test subsets. Then, a random forest algorithm was used to identify the variables predictive of a knee diagnosis group. Five different ML algorithms were developed, internally validated, and assessed by discrimination (area under the receiver operating characteristic curve, AUC), accuracy, precision (positive predictive value), recall (sensitivity), and F1 score (the harmonic mean of precision and recall).
Results: The models included patient characteristics and computer-assisted history taking. The support vector machine algorithm had the best performance for knee diagnosis group prediction, with good discrimination (area under the receiver operating characteristic curve, AUC = 0.92), accuracy (0.84), precision (0.85), recall (0.84) and F1-score (0.82).
Conclusions: The developed ML algorithm shows promise in predicting a knee diagnosis group in patients presenting with knee complaints to an orthopedic practice. Integrating this algorithm could streamline the consultation workflow by directing patients predicted to have knee osteoarthritis to orthopedic surgeons specializing in knee osteoarthritis, and those predicted to have ligamentous injuries to orthopedic surgeons specializing in sports and traumatic injuries.

☆ This article is part of a special issue entitled: 'AI and new technologies' published in The Knee.
* Corresponding author at: Jaffalaan 5, 2628 BX Delft, The Netherlands.
E-mail address: jacobienhfoosterhoff@gmail.com (J.H.F. Oosterhoff).

https://doi.org/10.1016/j.knee.2025.05.017
0968-0160/© 2025 The Author(s). Published by Elsevier B.V.

1. Introduction

The number of patients suffering from knee complaints is increasing [1], with an estimated prevalence of 365 million individuals annually affected by knee osteoarthritis (OA) [2,3], and an annual incidence of 70 per 100,000 individuals sustaining an anterior cruciate ligament injury [4]. Furthermore, general practitioners (GPs), who are usually the first point of contact for patients with knee complaints, face significant time constraints in their practice. This limits their ability to thoroughly evaluate musculoskeletal conditions before referring patients to an orthopedic surgeon. Moreover, orthopedic surgeons are becoming increasingly specialized, making it challenging for GPs to accurately refer patients to the most suitable specialist.

Using machine learning (ML) to predict a knee diagnosis group prior to an orthopedic consultation may be valuable in optimizing the consultation workflow. Patients likely to suffer from knee OA can be seen by an orthopedic surgeon who specializes in the diagnosis and treatment of OA, while patients who might have a ligamentous injury can have a consultation with an orthopedic surgeon who specializes in sports trauma.

Therefore, this study aimed to develop and internally validate an ML algorithm for predicting a knee diagnosis group for patients aged 18 years and older, based on computer-assisted history taking (CAHT), which includes patient characteristics and symptoms.

2. Materials & methods

This study was performed according to the Transparent Reporting of Multivariable Prediction Models for Individual Prognosis or Diagnosis Guideline (TRIPOD-Statement) (Supplementary Appendix 1) [5] and the JMIR Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research [6].

2.1. Ethical approval

This was a prospective, clinical cohort study using a convenience sample of patients 18 years of age or older, who visited our clinic for knee complaints from 1st May 2020, to 31st October 2020. This study was conducted in a single general district hospital and was deemed exempt from full review by the regional review board (study ID: N19.066).

2.2. Study design and population

In our hospital, all patients are referred electronically by their GP. Upon contact with the hospital, all patients 18 years or older referred for orthopedic care are asked to fill in a CAHT form in a secure online environment before consultation, for which they receive an e-mail with a hyperlink. In this CAHT form, patients indicate which joint they need to see an orthopedic surgeon for. Upon choosing a specific joint, the applicable follow up questions are presented. During the study period, all patients presenting with knee complaints as the reason for their consultation were identified. Following each scheduled consultation, the patients' electronic health records (EHRs) were reviewed to determine the recorded diagnosis. Each diagnosis was categorized into one of three groups: OA, ligamentous injuries, or other classifications. These categories were then added to a spreadsheet containing the patients' CAHT responses, creating an annotated dataset.

2.3. Primary outcome

The primary outcome of this study was the classification of knee diagnoses into three groups: knee OA, ligamentous injuries, and other classifications.

2.4. Candidate input variables

Based on clinical expertise, we selected potential predictive variables for model development. Next, a random forest (RF) algorithm with recursive selection was used to determine the feature importance for the outcome. RF algorithms for assessing variable importance are common in previously published orthopedic ML studies where variable importance was carried out [7]. The aim was to construct a model with variables relevant to clinical practice and useful for implementation.

2.5. Missing data

In the data pre-processing stage, patterns of missing data were assessed and determined to be missing by design. As a result, no missing data imputation was required.

2.6. Study population

In total, 1172 patients were included, with an average age of 54 years (interquartile range 36–66), of which the majority were female (*n* = 594, 50.7%) (Table 1). The most frequently occurring diagnosis was knee OA (*n* = 775, 66%), followed by ligamentous injuries (*n* = 208, 17.8%) and otherwise classified (*n* = 158, 13.5%).

2.7. Model development

First, the total dataset was divided into a training set (80%, *n* = 938) and a hold-out test set (20%, *n* = 234) stratified on the outcome (knee diagnosis group). Second, variables derived from the feature selection steps were used to train and internally validate various ML algorithms to choose the best-performing algorithm: penalized logistic regression (PLR), support vector machine (SVM), adaptive boosting (AdaBoost), decision tree (DT) and an RF algorithm. We trained each ML algorithm using five-fold cross-validation on the training set. The data were standardized before training the linear algorithms (PLR, SVM) using the *StandardScaler*. The data were not standardized before training the tree-based algorithms (AdaBoost, DT, RF), because these can inherently handle unstandardized data well.

2.8. Performance measures

The following measures were used for the assessment of predictive performance of the algorithms: discrimination (area under the receiver operating characteristic curve), balanced accuracy, precision (positive predictive value, PPV), recall (sensitivity), and F1-score (composite of precision and recall).

Discrimination was assessed with the area under the receiver operating characteristic curve (AUROC). The ROC curve plots the sensitivity (true positive rate) against 1 – specificity (false positive rate). The c-statistic ranges from 0.50 to 1.0 with 1.0 indicating the highest discriminating score and 0.50 indicating the lowest discriminating score. This differentiates between patients with knee OA, ligamentous injuries and other classifications [8].

Balanced accuracy is a metric used particularly when dealing with imbalanced datasets, calculated as the average of sensitivity (true positive rate) and specificity (true negative rate).

Precision, also known as PPV, refers to the proportion of patients with a positive prediction who actually had the outcome (TP/(TP + FP)). A perfect PPV is 100% but is highly dependent on the prevalence of the outcome; if the prevalence decreases, then the PPV decreases. Negative predictive value (NPV) refers to the proportion of patients with a negative prediction who did not have the outcome (TN/(TN + FN)). A perfect NPV is also 100%, but highly dependent on prevalence; if the prevalence decreases, then the NPV increases.

Recall, also known as sensitivity or true-positive rate, corresponds to the proportion of positive observations that are correctly classified as positive relative to all actual positive cases (TP/(FN + TP)).

F1-score tries to find a balance between precision and recall, and ranges from 0 (lowest) to 1 (highest).

In addition, the receiver operating characteristic curve and confusion matrix were visualized.

Table 1
Baseline characteristics of study population, n = 1172.

Variable	Median (IQR), n (%)
Age (years)	54 (36–66)
Female gender	594 (50.7)
Body mass index (kg/m^2)	26 (24–30)
Comorbidities	
Diabetes mellitus	57 (4.9)
Previous diagnosis cancer	84 (7.2)
Previous knee diagnosis	
Left knee	500 (42.7)
Right knee	530 (45.2)
Lower back	164 (14.0)
Previous operation	
Left knee	260 (22.2)
Right knee	273 (23.3)
Lower back	63 (5.4)
Diagnosis	
Osteoarthritis	775 (66.1)
Ligamentous injury	208 (17.7)
Other	189 (16.1)

IQR, interquartile range.

2.9. Statistical analysis

Data pre-processing and analysis were performed using Python 3.13. We used the following packages: pandas, numpy, sklearn, seaborn, and matplotlib.

3. Results

3.1. Variables importance

The RF identified the following variables as the most predictive for prediction of knee diagnosis category: age, trauma during sports activity, body mass index (BMI), duration of knee complaints, impact on sports or hobby activities, and pain at rest (Figure 1).

3.2. Performance of ML predicting knee diagnosis

The performance of the algorithms on the hold-out test set showed accuracies ranging from 0.78 to 0.84 (Table 2). Among them, the SVM algorithm achieved the highest performance, with an accuracy of 0.84 and an area under the curve (AUC) of 0.92, making it the best-performing algorithm and the final model of choice. Specifically, the SVM demonstrated strong discrimination in predicting knee OA with an AUC of 0.95 and ligamentous injuries with an AUC of 0.96 (Table 3, Figure 2).

The confusion matrix of the SVM algorithm showed that it accurately predicted 150 of 153 knee OA cases in the hold-out test set, resulting in a recall (sensitivity) of 0.98. For ligamentous injuries, the algorithm achieved a recall of 0.83, indicating reliable performance for both conditions (Table 3, Figure 3).

4. Discussion

We have developed and internally validated a prediction model using ML algorithms for predicting a knee diagnosis group in patients presenting with knee complaints to an orthopedic practice, achieving an accuracy of 84% and an overall AUC of 0.92. The prediction model is based on information retrieved from CAHT questionnaires, which included patient characteristics and symptoms. Integrating this algorithm could streamline the consultation workflow by directing patients pre-

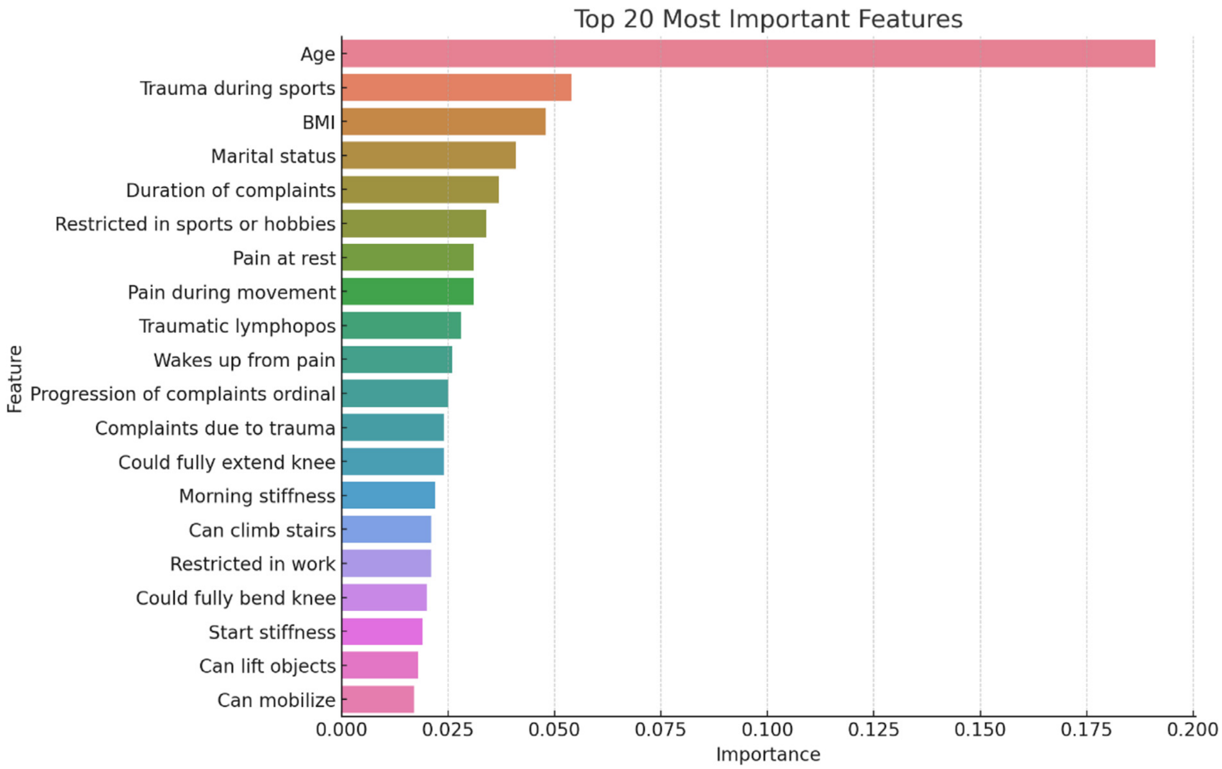

Figure 1. Variable importance random forest. BMI, body mass index.

Table 2
Model performances all algorithms in hold-out test set (n = 234).

Algorithm	Accuracy	AUC	Precision	Recall	F1-score
Penalized logistic regression	0.78 (0.76–0.80)	0.88 (0.84–0.91)	0.78 (0.74–0.83)	0.78 (0.76–0.80)	0.78 (0.75–0.81)
Support vector machine*	0.84 (0.79–0.88)	0.92 (0.89–0.96)	0.85 (0.81–0.89)	0.84 (0.80–0.89)	0.82 (0.75–0.87)
AdaBoost	0.80 (0.73–0.83)	0.83 (0.77–0.86)	0.80 (0.75–0.82)	0.80 (0.73–0.83)	0.80 (0.74–0.83)
Decision tree	0.78 (0.72–0.83)	0.77 (0.72–0.82)	0.78 (0.73–0.84)	0.78 (0.72–0.83)	0.78 (0.73–0.83)
Random forest	0.82 (0.77–0.87)	0.94 (0.91–0.97)	0.84 (0.78–0.89)	0.82 (0.77–0.87)	0.80 (0.73–0.85)

AUC, discrimination as measured by the area under the receiver operating characteristic curve.
* SVM chosen as the best-performing algorithm.

Table 3
Model performance support vector machine specified in hold-out test set (n = 234).

	Accuracy	AUC	Precision	Recall	F1-score
Balanced average	*0.84* (0.79–0.88)	*0.92* (0.89–0.96)	*0.85* (0.81–0.89)	*0.84* (0.80–0.89)	*0.82* (0.75–0.87)
Osteoarthritis	N/A	0.95 (0.92–0.98)	0.85 (0.80– 0.90)	0.98 (0.96–1.00)	0.91 (0.88–0.94)
Ligamentous injury	N/A	0.96 (0.93–0.98)	0.72 (0.58–0.86)	0.83 (0.69–0.94)	0.77 (0.66–0.87)
Other	N/A	0.86 (0.79–0.93)	0.94 (0.81–1.00)	0.37 (0.23–0.53)	0.53 (0.38–0.67)

AUC, discrimination as measured by the area under the receiver operating curve; N/A, not applicable.

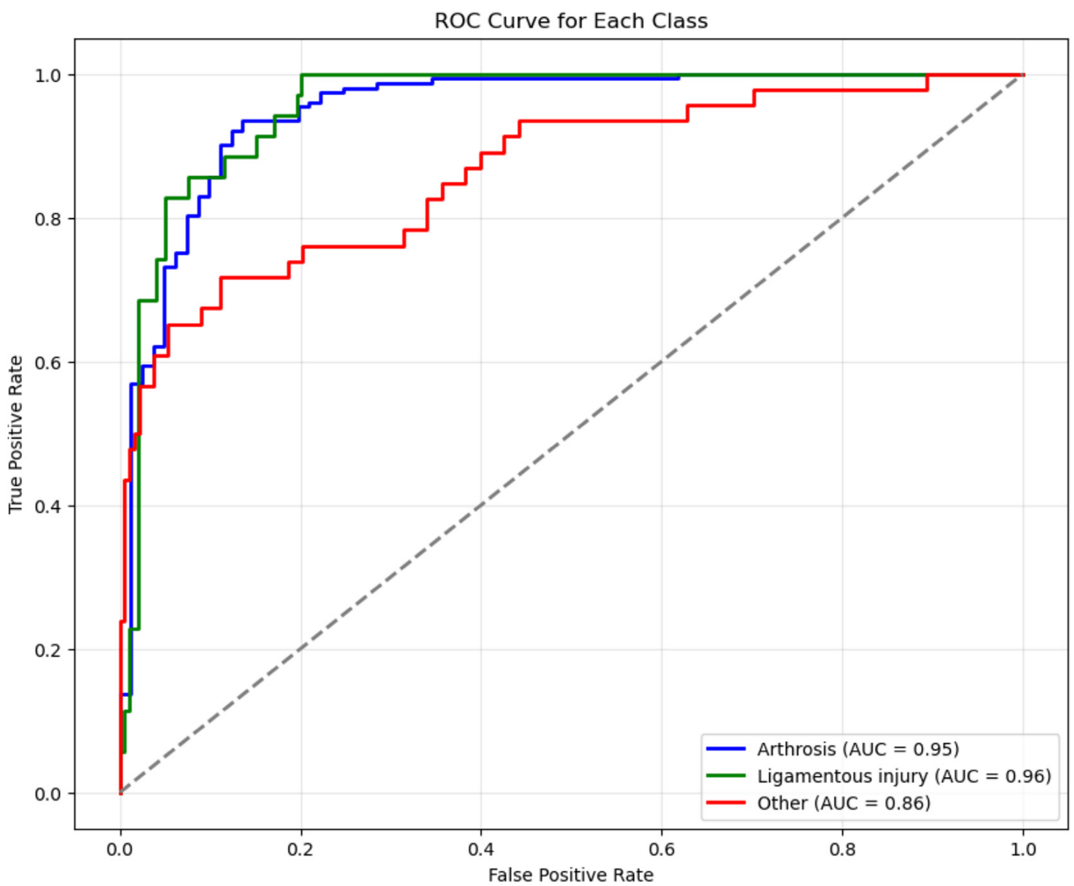

Figure 2. Receiver operating curves (ROCs) best-performing algorithm on the hold-out test set. AUC, area under the ROC curve.

dicted to have knee OA or other classified injuries to orthopedic surgeons specializing in conservative and operative management of OA, and those predicted to have ligamentous injuries to orthopedic surgeons specializing in sports and traumatic injuries.

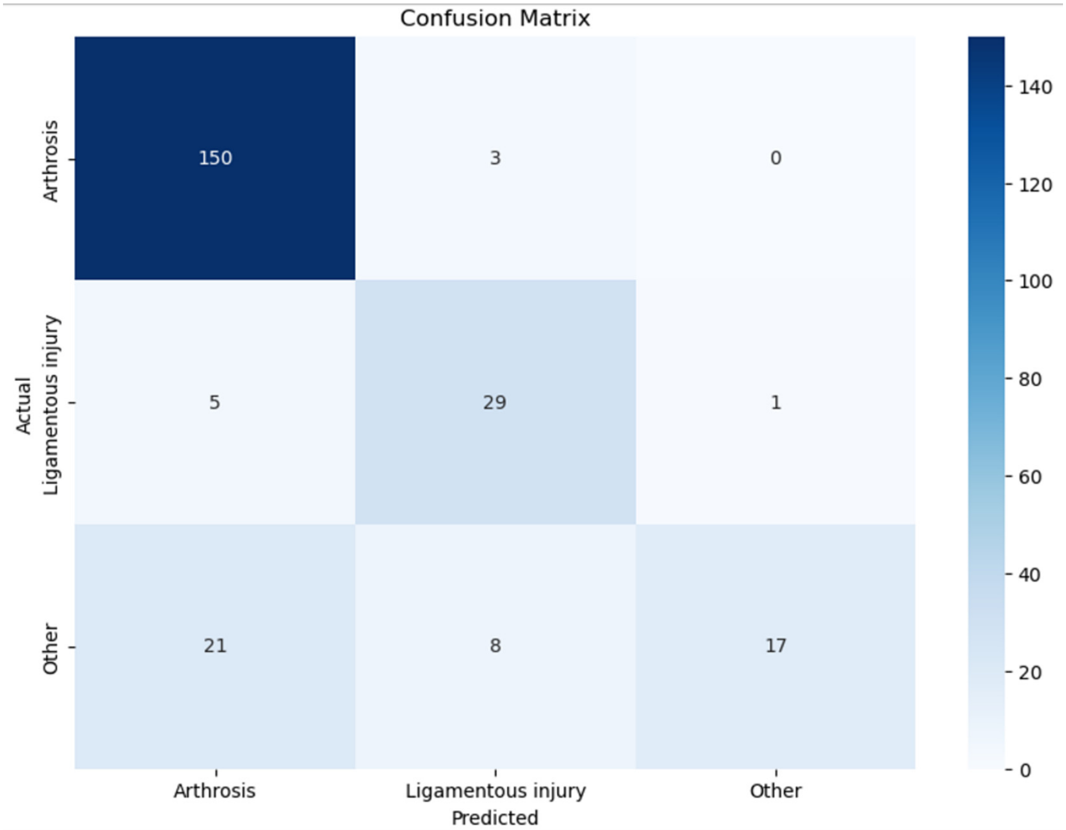

Figure 3. Confusion matrix best-performing algorithm on the hold-out test set.

4.1. Key findings

To the best of our knowledge, this is the first study applying ML algorithms to predict a knee diagnosis group in patients presenting with knee complaints before they are seen in the hospital. Numerous algorithms have demonstrated high accuracy in identifying knee pathologies; however, these primarily rely on radiographic or MRI imaging, which require a hospital visit [9,10]. An important subsequent step is moving from diagnostic to treatment and prognostic decision-making support. Musbahi et al. [11] demonstrated that ChatGPT (an open-access large language model) can make surgical decisions with confidence exceeding that of experienced knee surgeons. This suggests that integrating our ML algorithm that predicts a knee diagnosis group before the patient visits the hospital with an imaging-based knee pathology identification tool and then using a fine-tuned large language model for treatment decision support could substantially transform the current clinical workflow.

Our findings align with previous research. For example, van der Weegen et al. [12] recently investigated the use of CAHT in developing an ML algorithm to predict whether patients with hip complaints required operative or nonoperative treatment. Their approach has the potential to streamline consultation workflows by directing nonoperative patients to physician assistants and operative patients to orthopedic surgeons, including direct access to preoperative screening. This method aims to optimize the utilization of healthcare resources. The ML algorithm developed by van der Weegen et al. achieved an accuracy of 85%.

In our study, age, trauma, and BMI were identified as the most important features, which is in line with previous studies [13,14]. Marital status was considered a proxy variable, as we assume that the higher prevalence of OA in older patients, who often have a different marital status compared with younger patients, may account for its relevance.

The findings of our study suggest the potential to attenuate the logistical burden on musculoskeletal care (Figure 4). For example, this could be achieved by making this algorithm available to GPs, supporting them in their referral process to ensure that patients are seen by the most appropriate orthopedic surgeon for their knee complaints. With the rising burden of musculoskeletal injuries in primary care, GPs are increasingly tasked with triaging knee complaints, often without access to advanced diagnostic tools such as imaging. Given their gatekeeper role, accurate referrals are essential for a cost-effective healthcare system. By offering clinical decision support, this tool can assist GPs in distinguishing between frequently occurring pathologies – such as knee OA, ligamentous injuries, and other classified injuries – thereby streamlining referrals and

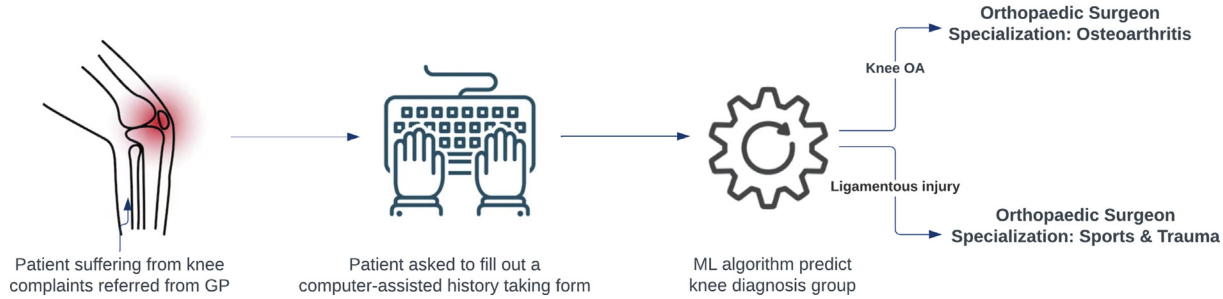

Figure 4. Process for machine-learning (ML) algorithm predicting a knee diagnosis group based on computer-assisted history taking (CAHT). GP, general practitioner; OA, osteoarthritis.

potentially reducing unnecessary consultations. While the achieved accuracy of 84% and AUC of 0.92 in our study are notable, differences in performance were observed across the knee diagnosis groups, with AUCs of 0.95 for knee OA, 0.95 for ligamentous injuries, and 0.86 for the otherwise classified group. The lower discrimination in the otherwise classified group may be attributed to its broad and heterogeneous nature. This means that false negatives may occur, leading to inaccurate subspecialty referrals and potential delays in care. Therefore, it is important to emphasize that the algorithm should be used as a clinical decision support tool, complementing the GP rather than replacing clinical judgment. Additionally, comparing the algorithm's performance to the accuracy of GP referrals to the correct subspecialty would help determine whether it meets or exceeds current clinical practice.

Future research will focus on refining predictions beyond broad knee diagnosis groups to improve accuracy for specific knee pathologies. For instance, knee OA could be further categorized based on the affected compartment (medial, lateral, or patellofemoral OA), while ligamentous injuries could be distinguished as medial and/or lateral collateral ligament injuries, posterior and/or anterior cruciate ligament injuries, or patellofemoral ligament injuries. Similarly, the broad "otherwise classified" category could be subdivided into clinically relevant diagnoses, such as meniscal injuries, tendon injuries, muscle injuries, bursitis, or other soft tissue conditions. These advancements aim to enhance the clinical utility and precision of the predictive model.

Although many promising ML algorithms exist, most remain in development or external validation phases, with only a few assessed in clinical practice [15,16]. We aim to prospectively evaluate our ML algorithm within our institution and validate it externally in an independent cohort. Because our algorithm is not classified as a medical device and is therefore exempt from medical device regulations [17], and is considered a low-risk AI system under the EU AI Act [18], these evaluations can proceed under current regulatory frameworks.

4.2. Limitations

The results of this study should be interpreted considering several limitations. First, the data were obtained from a single general district hospital in the Netherlands, which may limit its generalizability to international populations. Validation in independent hospitals would strengthen the reliability of the findings. Second, hyperparameter tuning was not performed, as the default settings of Python libraries were used. However, five-fold cross-validation was conducted on the training set to ensure model robustness. While hyperparameter tuning may have provided incremental benefits, we did not anticipate significant improvements in model performance given the characteristics of our cohort. Third, the development of an ML algorithm should focus on enhancing clinical care. While predicting knee diagnosis groups is valuable and may align with GPs' preliminary assessments, the maximum clinical utility could be achieved by extending predictions to operative versus non-operative management for patients with knee complaints [12]. Fourth, the current AUC of 0.92 indicates a strong ability of the model to distinguish between knee OA, ligamentous injury, and other injuries. The "other" injury group, which achieved an AUC of 0.86, presented greater classification challenges due to its heterogeneity. None the less, the model's performance for this group was still considered robust. Finally, the current study focused on the development of an algorithm and assessing performance, future research should also explore how patients perceive algorithm-assisted referral, including the impact on satisfaction with care.

5. Conclusions

In summary, the developed ML algorithm effectively predicted knee diagnosis groups for patients with knee complaints referred by GPs to orthopedic practices. The model has the potential to improve clinical workflow efficiency by guiding patients predicted to have knee OA to orthopedic surgeons specializing in OA treatment, and those with ligamentous injuries to surgeons specializing in sports and traumatic injuries. Future research should aim to further discriminate the group we have classified as the 'other' injury group. While this approach to allocating patients into diagnostic groups simplifies the complexity of clinical practice, advancements in healthcare-related artificial intelligence [19] are expected to significantly transform workflows in the near future.

CRediT authorship contribution statement

Jacobien H.F. Oosterhoff: Writing – original draft, Writing – review & editing, Methodology, Supervision, Investigation, Conceptualization. **Twan Slaats:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation. **Tristan Warren:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Walter van der Weegen:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization.

Funding

This research did not receive grants from any funding agency in the public, commercial or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge and thank the patients and staff contributing to the computer-assisted pre-hospital history taking acquisition.

References

[1] Cui A, Li H, Wang D, Zhong J, Chen Y, Lu H. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine 2020;29–30:100587. doi: https://doi.org/10.1016/j.eclinm.2020.100587.
[2] Long H, Liu Q, Yin H, Wang K, Diao N, Zhang Y, et al. Prevalence trends of site-specific osteoarthritis from 1990 to 2019: findings from the global burden of disease study 2019. Arthritis Rheumatol 2022;74:1172–83. doi: https://doi.org/10.1002/art.42089.
[3] World Health Organization. Osteoarthritis 2023. Available at: https://www.who.int/news-room/fact-sheets/detail/osteoarthritis [last Accessed June 2024].
[4] Paudel YR, Sommerfeldt M, Voaklander D. Increasing incidence of anterior cruciate ligament reconstruction: a 17-year population-based study. Knee Surg Sports Traumatol Arthrosc 2023;31:248–55. doi: https://doi.org/10.1007/s00167-022-07093-1.
[5] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med 2015;13:1–10. doi: https://doi.org/10.1186/s12916-014-0241-z.
[6] Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res 2016;18:e323. doi: https://doi.org/10.2196/jmir.5870.
[7] Dijkstra H, Oosterhoff JHF, van de Kuit A, IJpma FFA, Schwab JH, Poolman RW,, et al. Development of machine-learning algorithms for 90-day and one-year mortality prediction in the elderly with femoral neck fractures based on the HEALTH and FAITH trials. Bone Jt Open 2023;4:168–81. doi: https://doi.org/10.1302/2633-1462.43.BJO-2022-0162.R1.
[8] Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35:1925–31. doi: https://doi.org/10.1093/eurheartj/ehu207.
[9] Naguib SM, Saleh MK, Hamza HM, Hosny KM, Kassem MA. A new superfluity deep learning model for detecting knee osteoporosis and osteopenia in X-ray images. Sci Rep 2024;14:25434. doi: https://doi.org/10.1038/s41598-024-75549-0.
[10] Lee K-H, Lee R-W, Yun J-S, Kim M-S, Choi H-S. Automated diagnosis of knee osteoarthritis using ResNet101 on a DEEP:PHI: Leveraging a no-code AI platform for efficient and accurate medical image analysis. Diagnostics (Basel) 2024;14:2451. doi: https://doi.org/10.3390/diagnostics14212451.
[11] Musbahi O, Nurek M, Pouris K, Vella-Baldacchino M, Bottle A, Hing C, et al. Can ChatGPT make surgical decisions with confidence similar to experienced knee surgeons? Knee 2024;51:120–9. doi: https://doi.org/10.1016/j.knee.2024.08.015.
[12] van der Weegen W, Warren T, Das D, Agricola R, Timmers T, Siebelt M. Operative or nonoperative treatment is predicted accurately for patients who have hip complaints consulting an orthopedic surgeon using machine learning algorithms trained with prehospital acquired history-taking data. J Arthroplasty 2024;39:1173–1177.e6. doi: https://doi.org/10.1016/j.arth.2023.11.022.
[13] Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. Osteoarthr Cartil 2015;23:507–15. doi: https://doi.org/10.1016/j.joca.2014.11.019.
[14] Dong Y, Yan Y, Zhou J, Zhou Q, Wei H. Evidence on risk factors for knee osteoarthritis in middle-older aged: a systematic review and meta analysis. J Orthop Surg Res 2023;18:634. doi: https://doi.org/10.1186/s13018-023-04089-6.
[15] Dijkstra H, van de Kuit A, de Groot T, Canta O, Groot OQ, Oosterhoff JH, et al. Systematic review of machine-learning models in orthopaedic trauma. Bone Jt Open 2024;5:9–19. doi: https://doi.org/10.1302/2633-1462.51.BJO-2023-0095.R1.
[16] van de Sande D, van Genderen ME, Huiskens J, Gommers D, van Bommel J. Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit. Intensive Care Med 2021;47:750–60. doi: https://doi.org/10.1007/s00134-021-06446-7.
[17] The European Parliament and the Council of the European Union. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EE 2017. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32017R0745.
[18] The European Parliament and the Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139, 2024. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689.
[19] Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023;23:689. doi: https://doi.org/10.1186/s12909-023-04698-z.