EFORT open reviews

KNEE 7:7 491–497

The current role of the virtual elements of artificial intelligence in total knee arthroplasty

E Carlos Rodríguez-Merchán[1,2]

[1]Department of Orthopaedic Surgery, La Paz University Hospital, Madrid, Spain
[2]Osteoarticular Surgery Research, Hospital La Paz Institute for Health Research – IdiPAZ (La Paz University Hospital – Autonomous University of Madrid), Madrid, Spain

Correspondence should be addressed to E Carlos Rodríguez-Merchán
Email
ecrmerchan@hotmail.com

- The current applications of the virtual elements of artificial intelligence (AI), machine learning (ML), and deep learning (DL) in total knee arthroplasty (TKA) are diverse.
- ML can predict the length of stay (LOS) and costs before primary TKA, the risk of transfusion after primary TKA, postoperative dissatisfaction after TKA, the size of TKA components, and poorest outcomes. The prediction of distinct results with ML models applying specific data is already possible; nevertheless, the prediction of more complex results is still imprecise. Remote patient monitoring systems offer the ability to more completely assess the individuals experiencing TKA in terms of mobility and rehabilitation compliance.
- DL can accurately identify the presence of TKA, distinguish between specific arthroplasty designs, and identify and classify knee osteoarthritis as accurately as an orthopedic surgeon. DL allows for the detection of prosthetic loosening from radiographs.
- Regarding the architectures associated with DL, artificial neural networks (ANNs) and convolutional neural networks (CNNs), ANNs can predict LOS, inpatient charges, and discharge disposition prior to primary TKA and CNNs allow for differentiation between different implant types with near-perfect accuracy.

Keywords

- artificial intelligence
- total knee arthroplasty
- current role

EFORT Open Reviews
(2022) **7**, 491–497

Introduction

Artificial intelligence (AI) is an iterative process by which a machine captures information, transforms it into knowledge, and produces reactions that modify the environment. AI is a broad concept, involving virtual (computing) and physical (robotic) elements (1). AI is ubiquitous in today's society, as personal assistants (Alexa, Siri), viewing recommendation algorithms for video-on-demand platforms (Netflix), image and video processing applications (FaceApp), and self-driving cars (Tesla) (1). AI has enormous potential for improving health care; within a few years, it is feasible that AI will change the way daily clinical practice is conducted (2).

In this article, we have analyzed the virtual elements of AI with respect to their usefulness in total knee arthroplasty (TKA) (Table 1). A PubMed (MEDLINE) and Cochrane Library search of studies related to the virtual elements of AI in TKA was analyzed. The key words used were 'artificial intelligence AND TKA'. The main inclusion criteria were that the articles were focused on the virtual elements of AI. Studies not focused on such virtual elements were excluded. The searches were from the beginning of the search engines until 24 February 2022. The number of papers found was 93, of which 23 were finally chosen (Fig. 1).

What is artificial intelligence?

The term AI was established in 1956 by John McCarthy (3) who used this term to refer to computer capabilities and processes that are similar to human intelligence. AI implies a learning capacity, that is, the ability to perform tasks that have not been specifically programmed. An AI must be able to analyze information and make decisions in a similar way to a human being (4, 5).

Machine learning (ML) is a branch of AI that uses various systems and algorithms to learn and refine its operation through the use of data (6). Deep learning (DL) is a type of ML that can learn complex tasks by using large volumes of training information (7). DL uses an artificial neural network (ANN) composed of neurons arranged in a hierarchy. The network can process basic information at the initial level and forward it to the next level where it is integrated with data from other neurons and passed on to the next level. This process is performed iteratively until the system learns the task (e.g. analyzing, classifying, and segmenting radiological images) (8).

https://doi.org/10.1530/EOR-21-0107

Table 1 Virtual elements of artificial intelligence (AI).

Machine learning (ML)
Deep learning (DL)
*Artificial Neural Networks (ANNs)
*Convolutional Neural Networks (CNNs)

Convolutional neural networks (CNNs) are a DL subtype that are highly effective in image processing. CNNs use a complex series of layers and learnable filters through which data is passed, resulting in a final layer or output layer. When comparing CNNs to a conventional ANN used in DL, CNNs use the location of pixels in images to decrease the computational processing complexity and parameter requirements per layer (9).

One of the major advantages of DL and CNNs is their ability to learn independently which features in the input data lead to the results once the output data has been labeled. Given that CNN training is repetitive, there is a relationship between the database size and the algorithm's performance. DL algorithms require less time and computational power to analyze new data than other AI techniques (2).

Virtual elements of artificial intelligence (computer science) in TKA

This section reviews the current role and possible uses of ML and DL (ANNs and CNNs) in TKA.

Machine learning

In Table 2, the main ML models and their characteristics are summarized: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.

Validity in making predictions

Validity in predicting TKA component size According to Kunze *et al.*, predicting the size of prosthetic components for each patient before implanting TKA is crucial for avoiding the excessive costs associated with additional surgical trays and the morbidity associated with imperfect sizing (10). Kunze *et al.* demonstrated that ML algorithms can accurately predict the size of TKA components in real time. In their study, the authors analyzed 17 283 patients who underwent primary TKA between 2012 and 2020, with 9 different implant types. The primary variables were the final sizes of the femoral and tibial components extracted from automated inventory systems. Five ML algorithms were trained with routinely corrected demographic variables (age, height, weight, body mass index (BMI), and sex), using 80% of the study population and internally validated on an independent set of the remaining 20% of patients. The algorithm's performance was assessed by precision, mean absolute error, and root mean-squared error. The accuracy of the stochastic gradient boosting (SGB) model was 83.6% for predicting a ±4-mm deviation in the true anteroposterior femoral diameter and 95% for predicting a ±1 deviation in the size of the true femoral component. The SGB model's accuracy was 83% for predicting a ±4-mm deviation in the true medial/lateral tibial diameter and 97.8% for predicting a ±1 deviation in the size of the true tibial component. Patient sex was the characteristic that most influenced the SGB model's predictions for femoral and tibial component sizing. Subsequently, a TKA implant sizing application was created. The new ML algorithms published by Kunze *et al.* demonstrated good to excellent performance in predicting the size of TKA components. Patient sex played an important role in this prediction (10).

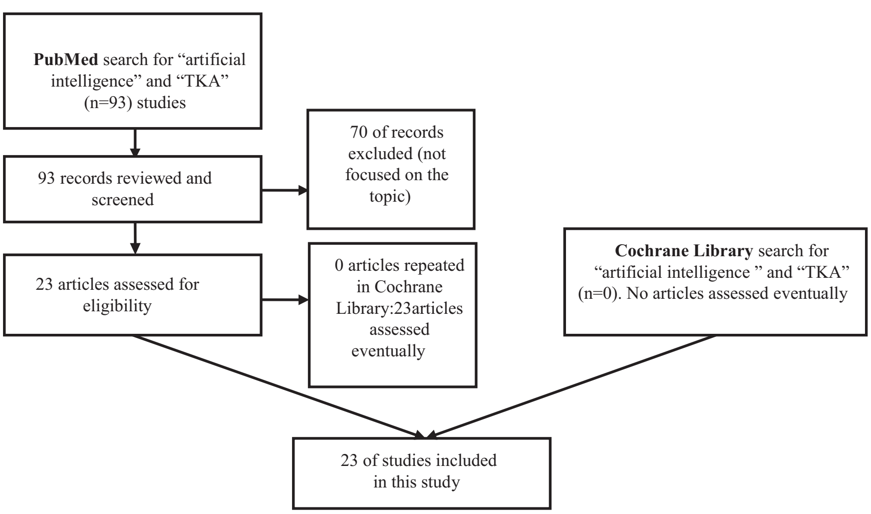

Figure 1
Flow chart of our search strategy regarding the role of artificial intelligence (AI) in total knee arthroplasty (TKA). The main inclusion criteria were that the articles were focused on the virtual elements of AI. Studies not focused on such virtual elements were excluded.

Table 2 Machine learning (ML) models.

Supervised learning	Unsupervised learning	Semi-supervised learning	Reinforcement learning
Data scientists supply input, output, and feedback to fabricate model (as the definition)	Uses deep learning (DL) to get conclusions and patterns through unlabeled training data	Fabricates a model through a combination of labeled and unlabeled data, a set of categories, suggestions, and exampled labels.	Self-interpreting but based on a system of recompenses and punishments learned through trial and error, looking for maximum reward.
Example algorithms			
*Linear regressions	*Apriori	*Generative adversarial networks	*Q-learning
+Risk evaluation	+Searcher		+Policy creation
+Sales forecasting	+Word associations	+Audio and video manipulation	+Consumption decrease
	+Sales functions	+Data creation	
*Support vector machines	*k-means clustering	*Self-trained Naïve Bayes classifier	*Model-based value estimation
+Image classification	+Performance monitoring	+Natural language processing	+Linear tasks
+Financial performance comparison	+Searcher intent		+Estimating parameters
*Decision tree			
+Predictive analytics			
+Pricing			

Validity in predicting length of stay and costs before primary TKA In 2018, Navarro *et al.* employed a predictive naïve Bayesian model to develop an ML algorithm using preoperative big data that could predict the length of stay (LOS) and inpatient costs after primary TKA (11). The study included 141 446 patients who underwent primary TKA. The algorithm's performance was calculated using the area under the receiver operating characteristic curve (AUC (area under the curve)) and the percent accuracy. The ML algorithm required several parameters, such as age, race, sex, and comorbidity scores ('risk of disease' and 'risk of morbidity'), demonstrating a high degree of validity, with an AUC of 0.7822 and 0.7382 for LOS and cost, respectively. As patient complexity increased, additional costs increased at levels of 3, 10, and 15% for moderate, major, and extreme mortality risks, respectively. This ML algorithm demonstrated excellent validity in predicting LOS and costs prior to primary TKA (11).

Validity in predicting transfusion after TKA Jo *et al.* conducted a level II evidence study to identify preoperative variables to create an ML model and to provide a web-based transfusion risk-assessment system for clinical use. The authors retrospectively reviewed 1686 patients who underwent TKA (12), collecting data on 43 preoperative variables, including medication history, laboratory values, and demographic characteristics. Variable selection was conducted using the recursive feature elimination algorithm. The transfusion group was defined as patients with hemoglobin levels <7 g/dL after TKA. A predictive model was developed using the gradient boosting machine, and the model performance was evaluated using the AUC. For external validation, data sets from an independent institution were tested with the model. Of the 1686 patients who underwent TKA, 108 (6.4%) were classified into the transfusion group. Six preoperative variables were selected, including preoperative hemoglobin level, platelet count, surgery type, tranexamic acid level, age, and body weight. The predictive model demonstrated good predictive performance using all six variables (AUC 0.842; 95% CI 0.820–0.856). Performance was also good according to external validation, using 400 data points from an independent institution (AUC 0.880; 95% CI 0.844–0.910). The web-based predictive model for transfusion after TKA using an ML algorithm with six preoperative variables was validated. The model proved to be simple and performed well, showing that it can be used prior to TKA to predict the transfusion risk, thereby enabling appropriate precautions to be taken for high-risk patients (12).

Validity in predicting patient dissatisfaction following primary TKA Kunze *et al.* conducted a study on 430 post-TKA patients to develop ML algorithms that could predict patient dissatisfaction (13). The authors performed a retrospective review of the patients between 2014 and 2016 and considered the following preoperative variables for the prediction: demographics, medical history, flexion contracture, knee flexion, and outcome scores (patient-reported health state, Knee Society Score (KSS) and KSS-Function). Recursive feature elimination was employed to select features that optimized algorithm performance. Five supervised ML algorithms were developed by training with ten-fold cross-validation three times. These algorithms were then applied to an independent testing set of patients and evaluated by discrimination, calibration, Brier score, and decision curve analysis. Forty (9%) patients were dissatisfied with the outcome after primary TKA after a minimum follow-up of 2 years. The random forest algorithm achieved the best performance in the independent testing set not used for algorithm development (c-statistic 0.77, calibration intercept 0.087, calibration slope 0.74, and Brier score 0.082). The most important predictors of dissatisfaction were age, number of medical comorbidities, presence of one or more drug allergies, preoperative patient-reported health state score, and preoperative KSS. This model demonstrated good discriminative ability in identifying patients at higher risk of dissatisfaction (13).

Validity in the identification of optimal sagittal component position in TKA In 2021, Farooq *et al.* attempted to find the optimal sagittal component position in TKA. Using ML algorithms, the authors performed a level III evidence study in which they identified the implant's sagittal position that could predict the best results (14). The authors retrospectively analyzed 1091 TKAs (67% female). All were posterior cruciate ligament-retaining or sacrificing with an anterior-lip (49.4%) or conforming-bearing (50.6%) and were performed with modern perioperative protocols. Preoperative and postoperative tibial slope and postoperative flexion of the femoral component were measured with standardized radiographic protocols. The analysis groups were classified according to satisfaction scores and the KSS question 'Does this knee feel normal to you?' ML algorithms were employed to identify the optimal sagittal alignment zones that predicted higher satisfaction scores and that resulted in the knee 'always feeling normal'. The mean age and median BMI were 66 years and 34 kg/m^2, respectively. The ML model predicted a higher likelihood of being 'satisfied or very satisfied' and that the knee 'always felt normal', with a tibial slope change closer to native (−2 to +2°) and a femoral component flexion of 0 to +7°. Poorer outcomes were predicted with any femoral component extension, femoral component flexion beyond +10°, and the addition or removal of >5° of the native tibial slope. That is, better patient-reported outcomes were predicted when approximation of the native tibial slope and incorporation of some flexion of the femoral component was achieved. Deviation from the native tibial slope and excessive femoral flexion or any extension of the femoral component predicted the poorest outcomes (14).

Validity of a wearable and machine learning-based surveillance platform

In a pilot study, Ramkumar *et al.* intended to validate the practicability of a remote patient monitoring (RPM) system in terms of the incidence of data interruptions and patient acceptance (15). Individuals downloaded the RPM mobile application before surgery to collect baseline activity and Patient-Reported Outcomes Measures (PROMs) data, and the wearable knee sleeve was paired to the smartphone during admission. The following parameters were collected up to 3 months after surgery: mobility (step count), range of motion, PROMs, opioid consumption, and home exercise program compliance. Validation was determined by the acquisition of continuous data and patient tolerance at semi-structured interviews 3 months after surgery. Of the 25 enrolled individuals, 100% had uninterrupted passive data collection. Of the 22 available for follow-up interviews, all encountered the system motivating and engaging. Ramkumar *et al.* stated that RPM offers the ability to more completely assess the patients experiencing TKA in terms of mobility and rehabilitation compliance (15).

Can machine-learning algorithms predict early revision TKA?

El-Galaly *et al.* studied whether ML algorithms predict early revision TKA in the Danish Knee Arthroplasty Registry (16). They used the aforementioned Registry to construct models to foresee the probability of revision TKA within 2 years of primary TKA. Age, post-fracture osteoarthritis, and weight were considered important preoperative factors within the ML models. During validation, the models' performance was not different from the non-informative models, and with area under the curves (AUCs) ranging from 0.57 to 0.60, no models reached the predetermined AUC threshold for a clinical useful discriminative capacity. Although a number of well-known presurgical risk factors for revision were coupled with four different ML methods, El-Galaly could not build up a clinically useful model capable of predicting early TKA revisions (16).

The prediction of distinct results with ML models applying specific data is already possible

In a systematic review with grade III of evidence, Hinterwimmer *et al.* analyzed ML algorithms for outcome prediction in TKA (17). The studies presented in such a review showed fair to good outcomes (AUC median 0.76/ range 0.57–0.98), while heterogeneous prediction models were analyzed: complications (6), costs (4), functional result (3), revision (2), satisfaction after surgery (2), surgical procedure (1), and biomechanical properties (1) were studied. The modified Coleman Methodology Score median was 65 (range 40–80) points. The conclusion was that the prediction of distinct results with ML models applying specific data is already possible; nevertheless, the prediction of more complex results is still imprecise.

Deep learning

Validity in detection and classification

Automated detection and classification of knee arthroplasty According to Yi *et al.*, preoperative identification of TKA is important for revision surgery planning. However, up to 10% of implants are not identified before surgery. The authors developed and tested the performance of a DL system to perform automated radiographic identification of the presence or absence of a TKA, differentiate a

TKA from a unicompartmental knee arthroplasty (UKA), and differentiate two different primary TKA models (18). The authors collected 237 anteroposterior knee radiographs with equal proportions of native, TKA, and UKA knees and 274 anteroposterior knee radiographs with equal proportions of 2 TKA models. Data augmentation was employed to increase the number of images for deep convolutional neural network (DCNN) training. A DL system based on DCNNs was trained on these images. Receiver operating characteristic curves with AUC were generated. Heatmaps were created using class activation mapping to identify the image features most important for DCNN decision-making. The two DCNNs trained to detect TKAs and distinguish between TKA and UKA achieved an AUC of 1. Heatmaps demonstrated appropriate emphasis of arthroplasty components in decision-making. The DCNN trained to distinguish between the two TKA models achieved an AUC of 1. Heatmaps showed an emphasis of specific unique features of the TKA model designs, such as the femoral component anterior flange shape. The DCNN was able to accurately identify the presence of a TKA and distinguish between specific arthroplasty designs (18).

Validity in automated detection and identification

Validity in identifying arthroplasty implants from knee radiographs According to Karnuta *et al.*, revisions and reoperations for patients undergoing TKA, UKA, and distal femoral replacement require accurate identification of the implant manufacturer and model. Failure to do so can delay medical care, increase morbidity, and increase the financial burden. The authors investigated whether a DL algorithm could accurately identify the TKA implant manufacturer and model from plain radiographs (19). They trained, validated, and externally tested a DL algorithm to classify TKA implants among nine different implant models from retrospectively collected anteroposterior plain radiographs. Performance was evaluated by calculating the AUC, sensitivity, specificity, and accuracy compared with a reference standard implant model from operative reports. The training and validation data sets consisted of 682 radiographs from 424 patients. After 1000 training epochs by the DL algorithm, the model discriminated nine implant models with an AUC of 0.99, an accuracy of 99%, a sensitivity of 95%, and a specificity of 99% on the external test data set of 74 radiographs. Ultimately, this DL algorithm, using single radiographs, differentiated nine TKA implants from four manufacturers with near-perfect accuracy (19).

Validity in preoperatively predicting value metrics for primary TKA Ramkumar *et al.* developed and validated an ANN that learns and predicts LOS, inpatient charges, and discharge disposition prior to primary TKA. They also applied the ANN to propose a risk-based, patient-specific payment model (PSPM) commensurate with case complexity (20). Utilizing data from 175 042 primary TKAs from the National Inpatient Sample and an institutional database, an ANN was created to predict LOS, charges, and disposition utilizing 15 preoperative variables. Outcome metrics included accuracy and AUC for a receiver operating characteristic curve. Model uncertainty was stratified by All Patient Refined comorbidity indices in establishing a risk-based PSPM. The dynamic model showed 'learning' in the first 30 training rounds with AUC of 74.8, 82.8, and 76.1% for LOS, charges, and discharge disposition, respectively. The PSPM showed that as patient comorbidity augmented, the risk increased by 2.0, 21.8, and 82.6% for moderate, major, and severe comorbidities, respectively. The reported DL model showed 'learning' with acceptable validity, reliability, and responsiveness in predicting value metrics, offering the ability to preoperatively plan for TKA episodes of care. This model may be applied to a PSPM proposing tiered reimbursements reflecting case complexity (20).

Artificial neural networks

Validity in making predictions

Artificial neural network prediction of same-day discharge following primary TKA based on preoperative and intraoperative variables In 2021, Wei *et al.* used an ANN model to determine the preoperative and perioperative variables that could predict same-day discharge for TKA patients (21). Data for their study were collected from the National Surgery Quality Improvement Program database from 2018. The authors included patients who underwent primary, elective, and unilateral TKA with the diagnosis of primary osteoarthritis. The authors analyzed demographic, preoperative, and intraoperative variables and compared the ANN model with a logistic regression model, which was a conventional ML algorithm. Variables collected from 28 742 patients were analyzed for their contribution to the length of hospital stay. The ANN model's predictability (AUC 0.801) was similar to that of the logistic regression model (AUC 0.796) and identified the following important predictors of same-day discharge: preoperative sodium level, international normalized ratio, BMI, age, type of anesthesia, operative time, dyspnea status, functional status, race, anemia status, and chronic obstructive pulmonary disease. Six of these variables were also statistically significant in the logistic regression analysis. In summary, both ANN modeling and logistic regression analysis revealed clinically significant factors for predicting the patients who can be safely discharged on the same day as the TKA procedure (21).

Table 3 Main current uses for artificial intelligence (AI) **systems** (virtual elements – computing) **in total knee arthroplasty** (TKA).

System used	Current uses of AI (reference)
Machine learning (ML)	• Prediction of the size of TKA components (10). • Prediction of the length of stay and costs before primary TKA (11). • Prediction of the risk of transfusion after primary TKA (12). • Prediction of the postoperative dissatisfaction after TKA (13). • Prediction of the poorest outcomes: Deviation from the native tibial slope and excessive femoral flexion or any extension of the femoral component can predict the poorest outcomes (14). • Validation of wearable and ML-based surveillance platforms (15). • Prediction of distinct results with ML models applying specific data (17).
Deep learning (DL)	• Accurate identification of the presence of TKA and differentiation between specific arthroplasty designs (18). • Differentiation between different implant types: A DL algorithm using plain radiographs can differentiate between nine TKA implants from four manufacturers with near perfect accuracy (19). • Prediction of value metrics for primary TKA (20).
Artificial neural networks (ANNs)	* Detection of perioperative factors that can predict discharge on the same day of surgery: preoperative sodium level, INR, BMI, age, type of anesthesia, operative time, dyspnea status, functional status, race, anaemia status, and COPD (21).
Convolutional neural networks (CNNs)	• Identification and classification of knee OA with the same precision as an orthopedic surgeon (22). • Detection of prosthetic loosening from radiographs. Its accuracy increases when using highly trained public algorithms and when clinical data are added to the algorithm (23).

COPD, chronic obstructive pulmonary disease; INR, international normalized ratio; OA, osteoarthritis.

Convolutional neural networks

CNNs can identify and classify knee osteoarthritis as accurately as a fellowship-trained orthopedic surgeon

According to Schwartz *et al.*, the use of a CNN to classify osteoarthritis severity could significantly reduce variability. The authors therefore conducted a study to retrospectively obtain knee radiographs from consecutive patients presenting for an arthroplasty consultation (22). The images were graded by four TKA surgeons using the International Knee Documentation Committee (IKDC) scoring system. The intraclass correlation coefficients (ICCs) for surgeons alone and for surgeons with a trained CNN were compared with 4755 different images. Four trained surgeons graded 1780 human knees. The ICC among the four surgeons for all possible IKDC grades was 0.703 (95% CI 0.667–0.737). The ICC for the four surgeons and the trained CNN was 0.685 (95% CI 0.65–0.719). The CNN was able to identify and classify knee osteoarthritis as accurately as an orthopedic surgeon (22).

Validity in the automated detection of implant loosening

In 2020, Shah *et al.* evaluated the ability of various CNN models to diagnose prosthetic loosening from preoperative radiographs and investigated the inputs that could improve their performance (23). The authors analyzed 697 patients operated on for first revision total hip arthroplasty or first revision TKA between 2012 and 2018. Preoperative anteroposterior and lateral radiographs and historical and comorbidity information were collected from the patients' electronic records. Based on the surgical notes, each patient was defined as having loose or fixed components. The authors trained a series of CNN models to predict the diagnosis of loosening. The constructed CNN provided good results in detecting loosening from radiographs alone. The first model built with only the radiographic images as input had an accuracy of 70%. The final model, which was built by fine-tuning a publicly available model named DenseNet, combining the anteroposterior and lateral radiographs and incorporating information from the patients' history, had an accuracy, sensitivity, and specificity of 88.3, 70.2, and 95.6%, respectively, on the independent test data set. The model's performance was better in the total hip arthroplasty revision cases (accuracy of 90.1%) than in the TKA revision cases (accuracy of 85.8%). This study demonstrated that CNN can detect prosthetic loosening from radiographs. The model's accuracy increased when highly trained public algorithms were employed and when clinical data were added to the algorithm (23). Table 3 summarises the main current uses of IA systems (virtual elements - informatics) in TKA.

Conclusions

The current use of the virtual elements (informatics) of IA in TKA offers several options of interest. ML can predict LOS and costs prior to a primary TKA; the risk of transfusion after a primary TKA; postoperative dissatisfaction after TKA; the size of TKA components; and poorer outcomes (deviation from the native tibial slope and excessive femoral flexion or any extension of the femoral component predict poorer outcomes). DL makes it possible to accurately identify the presence of TKA, distinguish between specific arthroplasty designs and identify and classify knee osteoarthritis as accurately as an orthopedic surgeon. ANNs can predict LOS, inpatient charges, and discharge disposition prior to primary TKA. CNNs can differentiate between different implant types with near-perfect accuracy and can detect prosthetic loosening from radiographs. A remote patient monitoring (RPM) system offers the ability to more completely assess the patients experiencing TKA in terms of mobility and rehabilitation compliance. The prediction of distinct results with ML models applying specific data is already possible; nevertheless, the prediction of more complex results is still imprecise.

ICMJE Conflict of Interest Statement
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding Statement
This work did not receive any specific grant from any funding agency in the public, commercial, or not-for-profit sector.

References

1. Anderson D. Artificial intelligence and applications in PM&R. *American Journal of Physical Medicine and Rehabilitation* 2019 **98** e128–e129. (https://doi.org/10.1097/PHM.0000000000001171)

2. Laur O & Wang B. Musculoskeletal trauma and artificial intelligence: current trends and projections. *Skeletal Radiology* 2022 **51** 257–269. (https://doi.org/10.1007/s00256-021-03824-6)

3. McCorduck P. Time line: the mechanization of thinking. In *Machines Who Think*, 2nd ed., pp. 23–30. Natick, MA: A. K. Peters, 2004. (https://doi.org/10.1201/9780429258985)

4. Liew C. The future of radiology augmented with artificial intelligence: a strategy for success. *European Journal of Radiology* 2018 **102** 152–156. (https://doi.org/10.1016/j.ejrad.2018.03.019)

5. Roman-Belmonte JM, De la Corte-Rodriguez H & Rodriguez-Merchan EC. Artificial Intelligence in Musculoskeletal Conditions . *Frontiers in Bioscience* 2021 **26** 1340 - 1348. (https://doi.org/10.52586/5027.PMID: 34856771)

6. Deo RC. Machine learning in medicine. *Circulation* 2015 **132** 1920–1930. (https://doi.org/10.1161/CIRCULATIONAHA.115.001593)

7. Suzuki K. Overview of deep learning in medical imaging. *Radiological Physics and Technology* 2017 **10** 257–273. (https://doi.org/10.1007/s12194-017-0406-5)

8. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H & Klang E. Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology* 2019 **290** 590–606. (https://doi.org/10.1148/radiol.2018180547)

9. Do S, Song KD & Chung JW. Basics of deep learning: a radiologist's guide to understanding published radiology articles on deep learning. *Korean Journal of Radiology* 2020 **21** 33–41. (https://doi.org/10.3348/kjr.2019.0312)

10. Kunze KN, Polce EM, Patel A, Courtney PM & Levine BR. Validation and performance of a machine-learning derived prediction guide for total knee arthroplasty component sizing. *Archives of Orthopaedic and Trauma Surgery* 2021 **141** 2235–2244. doi:10.1007/s00402-021-04041-5)

11. Navarro SM, Wang EY, Haeberle HS, Mont MA, Krebs VE, Patterson BM & Ramkumar PN. Machine learning and primary total knee arthroplasty: patient forecasting for a patient-specific payment model. *Journal of Arthroplasty* 2018 **33** 3617–3623. (https://doi.org/10.1016/j.arth.2018.08.028)

12. Jo C, Ko S, Shin WC, Han HS, Lee MC, Ko T & Ro DH. Transfusion after total knee arthroplasty can be predicted using the machine learning algorithm. *Knee Surgery, Sports Traumatology, Arthroscopy* 2020 **28** 1757–1764. (https://doi.org/10.1007/s00167-019-05602-3)

13. Kunze KN, Polce EM, Sadauskas AJ & Levine BR. Development of machine learning algorithms to predict patient dissatisfaction after primary total knee arthroplasty. *Journal of Arthroplasty* 2020 **35** 3117–3122. (https://doi.org/10.1016/j.arth.2020.05.061)

14. Farooq H, Deckard ER, Arnold NR & Meneghini RM. Machine learning algorithms identify optimal sagittal component position in total knee arthroplasty. *Journal of Arthroplasty* 2021 **36** S242–S249. (https://doi.org/10.1016/j.arth.2021.02.063)

15. Ramkumar PN, Haeberle HS, Ramanathan D, Cantrell WA, Navarro SM, Mont MA, Bloomfield M & Patterson BM. Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning-based surveillance platform. *Journal of Arthroplasty* 2019 **34** 2253–2259. (https://doi.org/10.1016/j.arth.2019.05.021)

16. El-Galaly A, Grazal C, Kappel A, Nielsen PT, Jensen SL & Forsberg JA. Can machine-learning algorithms predict early revision TKA in the Danish Knee Arthroplasty Registry? *Clinical Orthopaedics and Related Research* 2020 **478** 2088–2101. (https://doi.org/10.1097/CORR.0000000000001343)

17. Hinterwimmer F, Lazic I, Suren C, Hirschmann MT, Pohlig F, Rueckert D, Burgkart R & von Eisenhart-Rothe R. Machine learning in knee arthroplasty: specific data are key-a systematic review. *Knee Surgery, Sports Traumatology, Arthroscopy* 2022 **30** 376–388. (https://doi.org/10.1007/s00167-021-06848-6)

18. Yi PH, Wei J, Kim TK, Sair HI, Hui FK, Hager GD, Fritz J & Oni JK. Automated detection and classification of knee arthroplasty using deep learning. *Knee* 2020 **27** 535–542. (https://doi.org/10.1016/j.knee.2019.11.020)

19. Karnuta JM, Luu BC, Roth AL, Haeberle HS, Chen AF, Iorio R, Schaffer JL, Mont MA, Patterson BM, Krebs VE, *et al.* Artificial intelligence to identify arthroplasty implants from radiographs of the knee. *Journal of Arthroplasty* 2021 **36** 935–940. (https://doi.org/10.1016/j.arth.2020.10.021)

20. Ramkumar PN, Karnuta JM, Navarro SM, Haeberle HS, Scuderi GR, Mont MA, Krebs VE & Patterson BM. Deep learning preoperatively predicts value metrics for primary total knee arthroplasty: development and validation of an artificial neural network model. *Journal of Arthroplasty* 2019 **34** 2220.e1–2227.e1. (https://doi.org/10.1016/j.arth.2019.05.034)

21. Wei C, Quan T, Wang KY, Gu A, Fassihi SC, Kahlenberg CA, Malahias MA, Liu J, Thakkar S, Gonzalez Della Valle A, *et al.* Artificial neural network prediction of same-day discharge following primary total knee arthroplasty based on preoperative and intraoperative variables. *Bone and Joint Journal* 2021 **103-B** 1358–1366. (https://doi.org/10.1302/0301-620X.103B8.BJJ-2020-1013.R2)

22. Schwartz AJ, Clarke HD, Spangehl MJ, Bingham JS, Etzioni DA & Neville MR. Can a convolutional neural network classify knee osteoarthritis on plain radiographs as accurately as fellowship-trained knee arthroplasty surgeons? *Journal of Arthroplasty* 2020 **35** 2423–2428. (https://doi.org/10.1016/j.arth.2020.04.059)

23. Shah RF, Bini SA, Martinez AM, Pedoia V & Vail TP. Incremental inputs improve the automated detection of implant loosening using machine-learning algorithms. *Bone and Joint Journal* 2020 **102-B** (6_Supplement_A) 101–106. (https://doi.org/10.1302/0301-620X.102B6.BJJ-2019-1577.R1)