



# Artificial intelligence and machine learning as a viable solution for hip implant failure diagnosis—Review of literature and in vitro case study

Remya Ampadi Ramachandran<sup>1</sup> · Sheng-Wei Chi<sup>2</sup> · P. Srinivasa Pai<sup>3</sup> · Kharma Foucher<sup>4</sup> · Didem Ozevin<sup>2</sup> · Mathew T. Mathew<sup>1,5</sup>

Received: 24 August 2022 / Accepted: 9 January 2023 / Published online: 26 January 2023  
© International Federation for Medical and Biological Engineering 2023

## Abstract

The digital health industry is experiencing fast-paced research which can provide digital care programs and technologies to enhance the competence of healthcare delivery. Orthopedic literature also confirms the applicability of artificial intelligence (AI) and machine learning (ML) models to medical diagnosis and clinical decision-making. However, implant monitoring after primary surgery often happens with a wellness visit or when a patient complains about it. Neglecting implant design and other technical errors in this scenario, unmonitored circumstances, and lack of post-surgery monitoring may ultimately lead to the implant system's failure and leave us with the only option of high-risk revision surgery. Preventive maintenance seems to be a good choice to identify the onset of an irreversible prosthesis failure. Considering all these aspects for hip implant monitoring, this paper explores existing studies linking ML models and intelligent systems for hip implant diagnosis. This paper explores the feasibility of an alternative continuous monitoring technique for post-surgery implant monitoring backed by an in vitro ML case study. Tribocorrosion and acoustic emission (AE) data are considered based on their efficacy in determining irreversible alteration of implant material to prevent total failures. This study also facilitates the relevance of developing an artificially intelligent implant monitoring methodology that can function with daily patient activities and how it can influence the digital orthopedic diagnosis.

**Keywords** Artificial intelligence (AI) · Machine learning (ML) · Tribocorrosion · Acoustic emission (AE) · Random forest (RF)

## 1 Introduction

Recent medical diagnosis trends focus on identifying the significance and feasibility of intelligent models [1]. Literature details various studies in healthcare applications utilizing

technological advancements [2]. Our goal is to explore the field of data science, artificial intelligence (AI), and machine learning (ML), focusing on orthopedic implant diagnosis. Figure 1 gives some insight into the correlation between the above mentioned technological areas. As depicted in Fig. 1, ML is a subset of AI which includes a set of methods to detect patterns in data automatically. These uncovered patterns can assist in future data prediction and in performing effective decision-making. Key notion behind any ML model is learning from data. Data science is the interdisciplinary approach that comprises of mathematics, statistics, and computer science to effectively process these data to formulate actionable outcomes from it [3].

An orthopedic implant can fail due to its design flaws, manufacturing flaws, metallurgical failures, anti-bacterial, or anti-inflammatory reasons (infection), mechanical loosening, periprosthetic fractures, wear and osteolysis, dislocation, and surgical errors [4, 5]. Implant diagnosis is often performed when a patient complains about pain or discomfort, which is

✉ Remya Ampadi Ramachandran  
rampad2@uic.edu

<sup>1</sup> Department of Biomedical Engineering, University of Illinois Chicago, Chicago, IL, USA

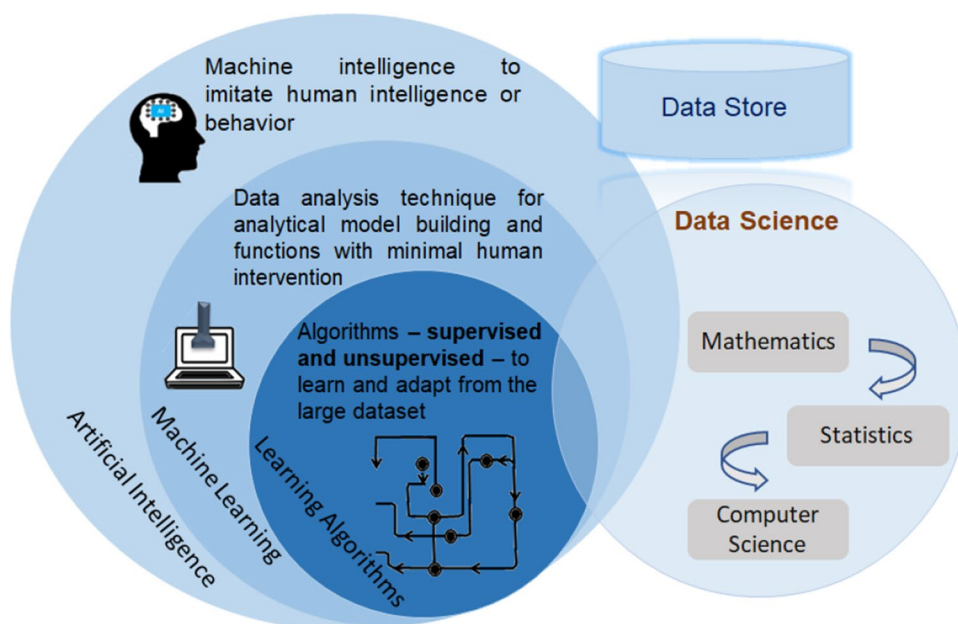
<sup>2</sup> Department of Civil, Materials, and Environmental Engineering, University of Illinois Chicago, Chicago, IL, USA

<sup>3</sup> Department of Mechanical Engineering, NMAM IT, Nitte, Karnataka, India

<sup>4</sup> Department of Kinesiology and Nutrition, University of Illinois Chicago, Chicago, IL, USA

<sup>5</sup> Department of Biomedical Sciences, University of Illinois at Rockford, Rockford, IL, USA

**Fig. 1** Schematic interpretation of the relation between data science, artificial intelligence (AI), and machine learning (ML)



often followed by a sequence of steps to confirm the problem [6]. Once an implant fails, revision surgery is the only available alternative with high risks involved. So, it is advisable to explore various options to identify the onset of failures and act as an early preventive measure to avoid any total or irreversible failure [7]. With the recent advances in the digital world, studies and trials with ML, AI, and data science are gaining popularity, and they are even considered for early diagnosis and preventive care in medical applications.

Thus, this paper is mainly driven by the following question: “Up to what extent the smart systems can find their future in orthopedic implant diagnosis?” This work is focused on exploring the applicability of ML, AI, and data science, through reported research studies, also investigating the need of a continuous monitoring tool that can predict the failure level of biomedical implants, especially those addressing THA implants.

## 2 Technological advancements in orthopedic diagnosis

As mentioned by Pei et al., digital orthopedics, along with advanced diagnostic imaging, surgical techniques, and advanced materials, supports surgeons in diagnosing and performing minimally invasive procedures to assist patient safety [8]. The usage of diagnostic imaging techniques is found to be promising in human health risk diagnosis but with challenges such as high operator-dependency (ultrasound), increased radiation burden (computed tomography CT), cost burden (magnetic resonance imaging MRI), and inability to predict local tissue reactions, presence of

tribocorrosion, degree of intra-operative soft-tissue destruction, and so on [9].

Orthopedic research focusing on joint replacement surgeries can be categorized based on the type of replacement procedure performed, such as hip, knee, ankle, wrist, shoulder, and elbow; implant failure category in sight; and diagnosis technique that is being used to identify the implant failure. The literature section has narrowed its focus to studies on orthopedic hip implant evaluation utilizing AI/ML or data science.

### 2.1 Literature review—Intelligent models for total hip arthroplasty (THA) monitoring

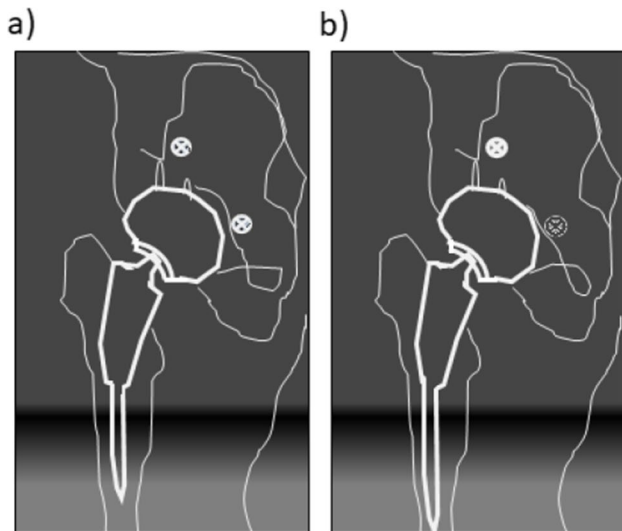
Present AI/ML studies for hip implants cover a wide range of applications, including implant detection, loosening detection, fracture classification, prediction of complications and wear rate, prediction of early revision surgery, predicting arthroplasty outcomes, and so on. Key aspects that need to be considered while designing an intelligent model are the selection of the dataset as well as the selection of appropriate algorithms for the dataset to provide better performance or accurate outcomes. These ML studies are grouped in this section based on the model’s outcome, as depicted in Table 1. It is also noted that radiographs are the most commonly used dataset for ML hip-related studies.

#### 2.1.1 Implant detection models

The primary aim of implant detection models is to confirm the presence or absence of implants from diagnostic images (Fig. 2). If developed considering appropriate training

**Table 1** Roadmap of existing ML algorithms deployed for monitoring or performing decision-making after THR surgery is categorized as depicted based on their applications

AI/ML models	
Implant detection models	Implant failure prediction models
Post-operative decision-making models	Pre-operative decision-making models
Clinical and patient-reported outcomes models	Structural monitoring models
Simulation models	



**Fig. 2** Example of radiographs (drawings) that can be given as input to the ML model for implant detection

factors, these models can play a vital role in clinical decision-making [10]. One of the models that work on YOLOv3 convolutional neural network (CNN)-based approach can function as a stem detection model using the postoperative hip anteroposterior (AP) X-rays dataset [11]. Karnuta et al. explored the possibility of developing an artificially intelligent model to identify the implant manufacturer and model that learn based on AP plain radiographs and perform decision-making. This scheme also utilizes CNN models and supports pre-operative analysis before revision surgeries but can be applied only for the femoral component [12]. Another deep CNN approach described in [13] supports pre-operative planning before revision surgery implants and risk prevention. This model relies on AP hip X-ray dataset to generate a fully automatic and interpretable model that can identify the design of THA. The authors also claim that it is a novel study in this field that can be advanced to an AI method incorporating supplementary radiographic views. Another approach confirms that pre-operative identification of failed hip implants could help properly plan revision surgeries [14]. ResNets and k-nearest neighbors (k-NN) achieve anatomical localization and classification of metallic implants, especially implants in the femur, with whole-body post-mortem computed tomography (PMCT) scans as the dataset, which

can help to draw pattern-recognition and conclusions in a clinical and forensic setting [15].

Deployment of the implant detection ML model is effective in implant diagnosis and performing properly planned decision-making before revision surgery. It can also assist in identifying the design and manufacturer of the failed implants where a manual evaluation is challenging.

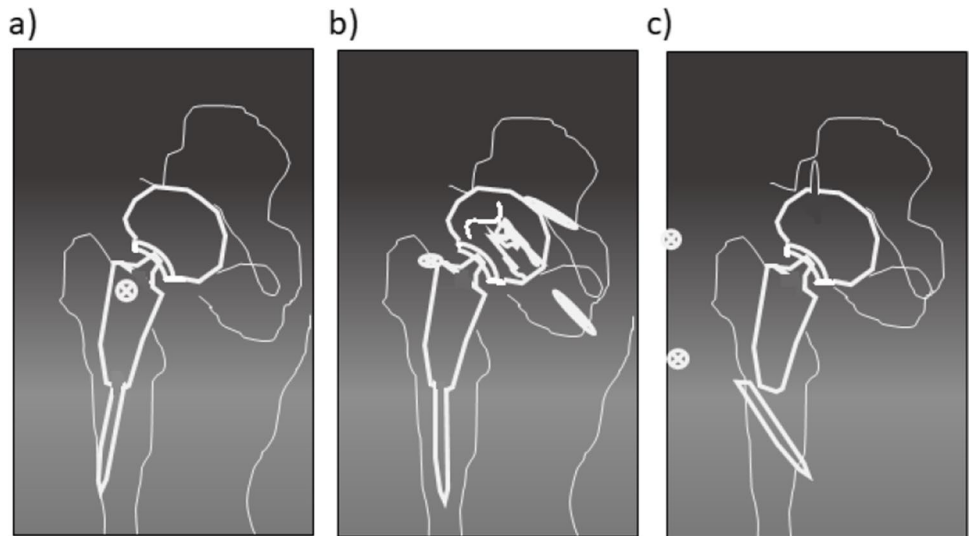
### 2.1.2 Implant failure prediction models

Hip implant failure can be due to its loosening from bone caused by wear and tear, or tribocorrosion, failed osseointegration, infection, damaged bearings causing malposition of its parts, and crack or fracture. ML studies dealing with such failures can act as secondary evidence in clinical decision-making by bridging the gap between evidence-based medicine and the patient’s personal context [10]. Pre-operative AP and lateral radiographs with the CNN model predict hip implant loosening with an accuracy of 90.1%, concluding that ML models can predict implant loosening using radiographs [16]. Another use of ML is observed when it can predict hip fractures from the radiographs with an increased accuracy of 19% compared to human evaluation. This model also utilizes a CNN-based localization and classification approach using AP pelvic and AP radiographs [17]. Since the malposition of the acetabular component relies on pelvic-sagittal inclination, its assessment can support patient-specific THR surgery planning. Jodeiri et al. developed a fully automated scheme that can measure the inclination of THR and assist in clinical decision-making using CT images [18] (Fig. 3).

The simultaneous wear and corrosion in the biological environment can cause implant failures. Though studies relating to this viewpoint are limited, the corrosion severity estimation model at the stem taper of retrieved THA implants employs a pattern-recognition approach for the image dataset [19]. This model with SVM Bayesian optimization gave an accuracy of 85% for the visual scoring method. Rouzrokh et al. developed a model to assess the risk of hip dislocation [20] and a fully automated tool that measures the acetabular component angles from the radiographs [21]. Based on the results, this tool can interpret the risk of hip dislocation from radiographs.

Even though implant failure prediction is highly significant, dictated ML models mainly consider radiographic

**Fig. 3** Understanding implant failures from a radiograph (drawings): **a** loosening, **b** malposition due to loosening, **c** fracture; examples of ML model inputs



images as their main source of learning data. It limits the chances of predicting failures before it happens and opens the door for ML models that can even predict the chances of failures.

The following sections from Sects. 2.1.3 to 2.1.5 discuss ML models that consider patient’s clinical or electronic records as their main learning source. Any failure to comply with routine follow-up can lead to data deficiency affecting the model’s accuracy, which highlights the need for data sets of these types of models needs to be updated regularly for a promised outcome.

**2.1.3 Postoperative decision-making models**

Here the main focus of ML models is to predict the chances of revision surgery after primary THR. Patient characteristics such as clinical variables, demographics, comorbidities,

cognitive appraisal processes, and surgical variables are considered following revision surgery indications such as aseptic loosening, dislocation/instability, periprosthetic joint infection, periprosthetic fracture, adverse local tissue reaction/metallosis, and miscellaneous (Table 2). Based on these patient factors, Kunze et al. developed a clinical decision-making tool to confirm its discriminative capability in performing patient health assessment after THA [22], while Klemt et al. confirmed the potential of the ML models to support clinical practice in quantifying the risk of revision THR surgery through patient-specific characteristics [23]. Another different approach that predicts postoperative THA outcome score at 3 months utilizes the least absolute shrinkage selection operator (LASSO) for its ML model. In this work, Sniderman et al. confirmed that cognitive appraisal processes contribute more to determining the postoperative hip disability and osteoarthritis outcome score (HOOS) [24].

**Table 2** Sample patient medical record

Sample Medical Record: Name	
Clinic:	
Patient Information	
Problems	Review of Systems
Medications	Vital Signs
Directives	Physical Exam
Allergies and Adverse Reactions (! = critical)	Plan
Service Due	Medications:
OFFICE VISIT	Treatment:
History of Present Illness	Orders:
Reason for visit: Routine to follow up	Education/Counseling (time):
Chief Complaint: No complaints	Coordination of Care (time):
	Follow-up/Return Visit:
History	Disposition:
Symptoms	Tests:

### 2.1.4 Pre-operative decision-making models

Predicting complications after primary THA can help in pre-operative decision-making. Mortality rate prediction is found to be possible by deploying artificial neural network (ANN) and logistic regression (LR) models. One of the models considers elderly patients with THA implants after hip fractures as their study subject [25]. This model gives insights into clinical outcomes and the likelihood of mortality, suggesting a pre-operative decision-making approach for doctors, patients, and even their families. A pre-operative ML model developed by Karhade et al. predicted opioid use after THA [26]. While American Joint Replacement Registry (AJRR) risk calculator model shows poor performance or discrimination for 90-day mortality samples [27], other studies from the same team show improved accuracy in predicting 30-day mortality and cardiac outcomes using LASSO regression models in Veterans Health Administration patients [28] and American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) [29] data.

However, an early risk stratification study using the ACS-NSQIP universal risk calculator by Edelstein et al. facilitates an online post-surgery complication prediction model [30]. Also, in [31], Shah et al. discussed the possibility of developing an ML-based prognostic model to determine the peri-operative risks, which can help identify and address potential risk factors leading to various complications. Implementing ML models in predicting patients’ ambulatory same-day discharge after THA seems interesting due

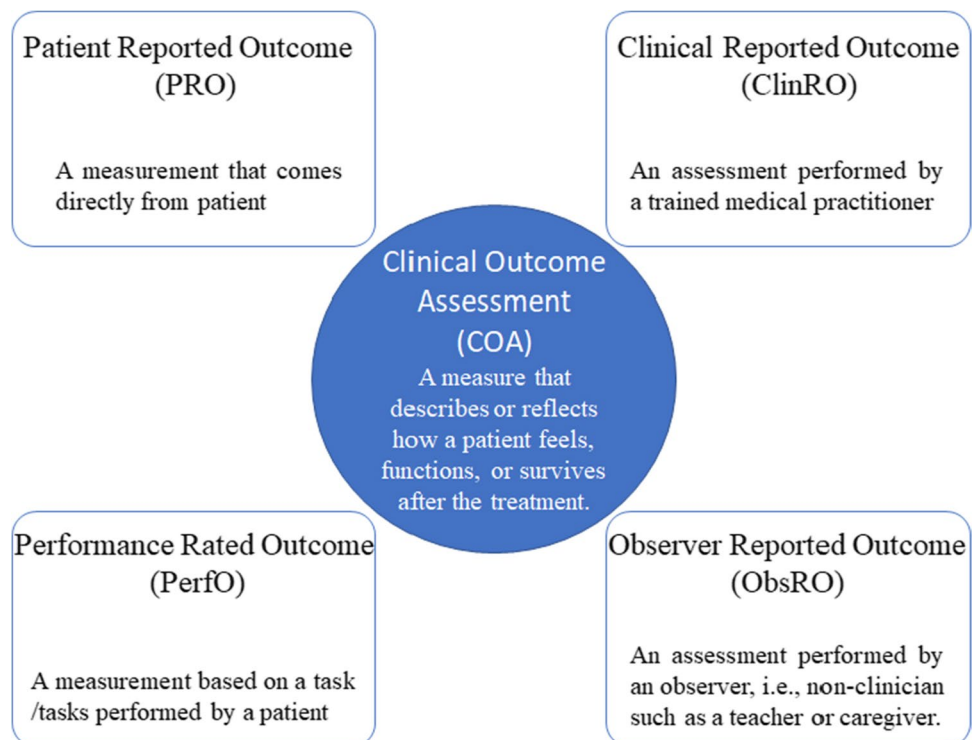
to continuing financial and regulatory pressures [32]. This study also confirmed the patient candidacy based on clinical variables, demographics, and other comorbidities (Table 2).

### 2.1.5 Clinical and patient-reported outcomes models

The K-means clustering ML method based on gait or sensor data and electronic data records on peri-operative setup is useful in predicting clinical or patient-reported outcomes measures (PROM) of TJA surgeries [33]. It also helps cluster patients parallel to the predicted PROMs with an opportunity for both qualitative and quantitative features in model development (Fig. 4). Another PROM-based ML model used the logistic LASSO model and HOOS scores to determine the minimally clinically significant difference (MCID) and facilitate presurgical patient education and decision-making [34].

Risk assessment of patients with THA is typically performed subjectively using questionnaires or gait analysis. Here, supervised classifier models such as support vector machine (SVM) and linear discriminant analysis (LDA) on sensor-derived metrics (for the gait analysis) show improved prediction accuracy after their primary THA [35]. An administrative big data analysis model based on the naïve Bayesian ML algorithm with demographics and comorbidities dataset supports the prediction of length of stay and payment before primary THR surgery [36]. Altogether, models categorized in this group focus on either the clinical or reported outcomes after replacement surgery.

**Fig. 4** Treatment outcome including patient reported and clinical-reported outcomes

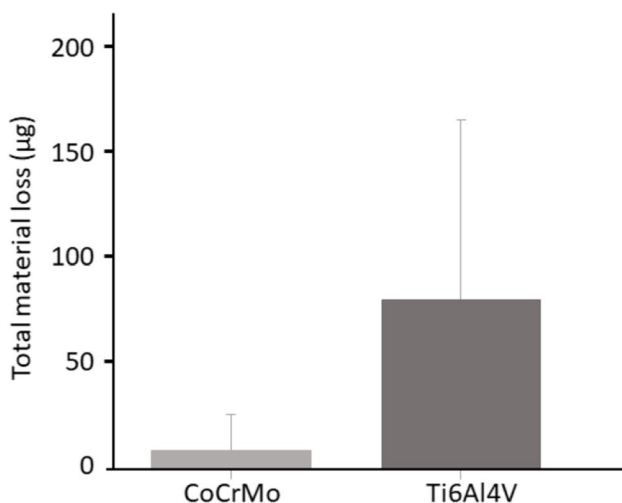


### 2.1.6 Structural monitoring models

Most of the studies described so far are focused on patient experiences after the THR surgery and monitoring is purely based on radiographs, image-based analysis, or electronic databases. Attention to structural monitoring and quantitative analysis is lacking here, which can determine the presence of local tissue reactions, tribocorrosion, and other-related failures. Even though Borjali et al. reported an ML model that can predict polyethylene wear rate in pin-on-disk of hip implants, a lack of continuous assessment is evident from its learning data [37]. Based on the prediction error of ML algorithms, k-nearest neighbor (KNN) was identified as the best predictive model for the selected dataset, which is purely based on the previous literature dataset. In order to generate a wear rate dataset, manual intervention is necessary as it lacks continuous monitoring of implants. From the alloys selected for this study, even though Ti alloy is an acceptable clinical alloy, it has inferior wear resistance compared to CoCrMo alloy. On the other hand, CoCrMo is more toxic than Ti alloy [38, 39]. Figure 5 is an example of the wear rate observed during a tribocorrosion experiment for hip damage assessment of Ti6Al4V and CoCrMo alloys [40].

### 2.1.7 Simulation models

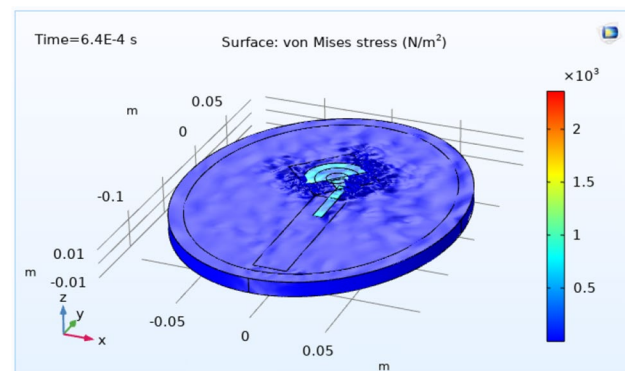
The advantages of simulation models mainly reside in developing a digital prototype. It enables creating and analyzing complex problems along with safety, efficiency, repeatability, and cost-effectiveness moderately. It can be a predictive model to validate the results when processes or external factors change. Simulation studies often function based on



**Fig. 5** Ti6Al4V exhibits higher total material loss due to wear and corrosion ( $K_{WC}$ ) in micrograms ( $\mu\text{g}$ ) than CoCrMo

the finite element method (FEM), a numerical technique to replicate any physical phenomena [41].

A list of simulation studies is reported in this literature, discussing studies from an ML aspect. Al-Dirini et al. performed a feasibility study to check the implant stability of cemented femoral stem. This study includes a combination of finite element (FE) models and surrogate Gaussian models for implant micromotion assessment and confirms its applicability to determine possible implant position variations and assist in the early THA implant design cycle [42]. In order to understand the quality of life in osteoarthritis patients, Kreif et al. used a simulation model with relative bias, efficiency, and confidence interval coverage of the selected methods, incorporating machine learning and misspecified parametric models [43]. Custom-made hip prosthesis method discussed in [44] details HIP prostheses manufacturing innovative COMputing (HIPCOM) tool and the HIPCOM design environment (HIDE) that can support enhanced effective prosthetic design for the patient to improve the reconstructive surgery. Using FE analysis, ML classification, and regression models, Ricciardi C. et al. developed prosthetic decision-making and long-term surgery-outcome model [45]. The applicability of deep learning-based CNN models for constructing 3D anatomy from 2D x-ray images discussed by Almeida et al. can help in surgery planning and diagnosis [46]. Another ML technique presented in [47] explored the possibility of FE analysis and ML algorithms to evaluate further optimization of the short-stem hip prosthesis to reduce the stress shielding effect and to deliver better performance, where stress shielding, otherwise termed as stress protection, refers to the reduction in bone density causing (mechanical) bone loss. Figure 6 is a representative simulation model for the acoustic emission (AE) monitoring study using the COMSOL Multiphysics simulation tool [48]. This model is developed to predict the impact of soft-tissue layers on stress wave or AE signal propagation initiated due to implant failures or defects.



**Fig. 6** Example of COMSOL simulation model representing stress wave simulation due to loading activities experienced by THR

## 2.2 Discussion: Intelligent models for total hip arthroplasty monitoring

Many AI–ML studies use radiographs and electronic data documents as their learning source for identification and prediction. The ultimate use of ML models in radiology research as well as other biomedical applications might be due to ML algorithm's capability to recognize complex patterns in the images or data to make intelligent decisions. Based on the training data, ML models can identify precise patterns and deliver accurate outcomes. Most importantly, when an ML model is being developed, it is a good practice to incorporate almost every possible characteristic or feature the model needs to be trained on. Also, training data should be preprocessed to avoid any misleading dataset during its learning phase. An ML model developed in this manner will be able to handle most of the scenarios and complex patterns, thereby being able to generalize the knowledge (like human intelligence) to deliver an accurate outcome for an unknown dataset [49, 50]. Even with the highest accuracy outcomes, our perspective is that these studies lack a continuous implant monitoring possibility. Since there are many reasons for implant failures that can eventually lead to revision surgery [51], a technique that can identify and predict the onset of failure would be advantageous for early diagnosis and adopting preventive measures.

Table 3 illustrates the existing AI/ML studies related to THR, learning models, and the dataset adopted for each study. When looking at this table from a routine post-surgery monitoring perspective, some questions may arise (1) *Is it possible to feed the model with radiographs on a routine basis?* (2) *Are we updating the electronic database or patient records on a regular basis once the surgery is over and the patient is discharged?* Otherwise, the chances of updating these databases will happen only when the patient comes back with some complaints or for any wellness visit, which is counted as one of the major limitations of existing procedures. Also, most datasets discussed so far have limitations, such as operator dependency, radiation and cost burden, inability to predict local tissue reactions, corrosion, soft-tissue destruction, etc. [9].

The comparison chart reveals a lack of ML models focusing on routine post-surgery monitoring to facilitate preventive care and avoid high-risk revision surgery. One of the common implant failure reasons that come into the picture is tribocorrosion which is an irreversible alteration of implant material, especially at the interfaces. Since implant tribocorrosion can eventually lead to many health-related issues, including toxicity, crack, and fracture, it is necessary to highlight the relevance of a scheme that can detect and predict tribocorrosion failures and function as a non-invasive routine post-surgery monitoring scheme. This highlights the relevance of research [40, 52], reporting AE's effectiveness

in predicting bio-tribocorrosion in dental and orthopedic implants facilitating continuous monitoring opportunities. Because AE is exclusively focused on active defects, any tribocorrosion event on the implant interface can initiate AE signals. These tribocorrosion events can eventually cause structural damage and permanent failure of the installed implant. Previous study results proved that tribocorrosion assessment is possible with AE technique. The comparison between selected features also indicates that higher absolute energy, higher friction coefficient, and lower corrosion potential are indications of increased corrosion [40]. A computational model study reveals the impact of tissue thickness in AE signal transmission and its detectability at the skin surface or AE receiver sensor [53].

Based on the findings in [40, 52, 53] and the AI/ML model's capability to handle a large volume of data, further investigations on the development of THR structural monitoring tools are in progress. Once a model is developed with a large volume and high-quality learning data, it can provide accurate interpretations with minimal or no errors and minimal human intervention in decision-making. A recent study [54] shows the possibility of post-market medical device surveillance using electronic health records (EHR) data. However, we believe that surveillance of prostheses once implanted might not be complete with EHR data unless a continuous routine monitoring system along with EHR is in place to report any medical anomalies after surgery.

Nich et al. [55] discussed the applicability of AI and ML for hip and knee surgery but its applicability in clinical practice is still overdue. Even though AI/ML studies reported in the literature are capable of predicting implant loosening [16], hip fracture [17], angular position assessment [20, 21], these studies solely relied on radiographs extending an expert dependency. Likewise, most ML models reported in the literature use radiographs, CT scans, X-rays, and electronic data records as inputs or main learning data sources. Other studies relevant to this context include predicting polyethylene wear rate using the data collected from the literature [37] and corrosion severity estimation from stem taper images [19]. It is evident that a continuous implant monitoring feasibility study is relevant. We also believe that the factors that lead to implant failure include mechanical loosening, periprosthetic fracture, implant fracture, infection, and wear/corrosion/osteolysis/tribocorrosion, as shown in Fig. 7 [5]. Identifying and predicting tribocorrosion damage in an implant is significant to take preventive measures and avoid failures, ultimately leading to a complication or a revision surgery [40, 53].

Reviewing existing ML models in hip implant diagnosis, its benefits, and limitations gives an insight into the state-of-the-art and knowledge gap in this field. As part of a comparative study, we did deep dive into a variety of ML studies in THR diagnosis and could see that image-based models

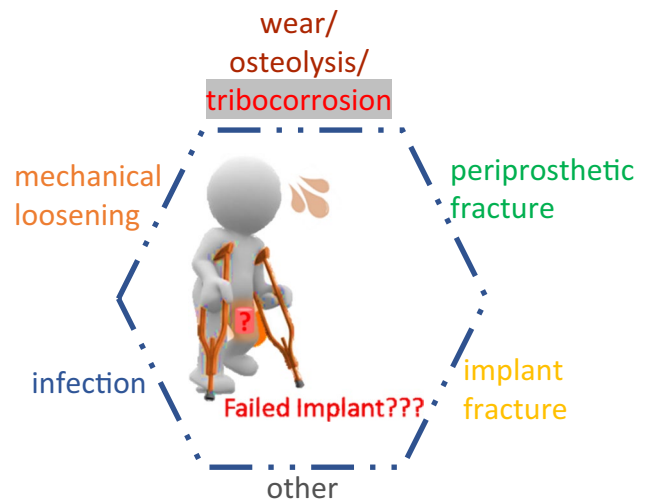
**Table 3** Summary of existing AI/ML models related to THR procedures

ML algorithm	Dataset	Outcome	Reference
<i>Detecting implants for analyses after THR surgery</i>			
YOLOv3 CNN	Postoperative hip anteroposterior (AP) X-rays	Stem detection model	Kang et al. [11]
CNN models	AP plain radiographs	To identify hip arthroplasty implants	Karnuta et al. [12]
Deep CNN	AP hip X-ray	Identify the design of THA implants	Borjali et al. [13]
NN	Postoperative anteroposterior (AP) hip radiographs	Identification and classification of femoral stem	M. Murphy et al. [14]
<i>Failure analyses after THR surgery</i>			
CNN model	Pre-operative AP and lateral radiographs	Predicts hip implant loosening	Shah et al. [16]
CNN-based localization and classification	AP pelvic radiographs and AP radiographs	Predict the hip fractures	E. A. Murphy et al. [17]
SVM Bayesian optimization	Stem taper images	Corrosion severity estimation	Milimonfared et al. [19]
U-Net CNN model	Anteroposterior (AP) pelvis, cross-table lateral hip postoperative radiographs	Automated measuring tool for acetabular component angular position	Rouzrokh et al. [21]
<i>Post-operative decision-making</i>			
SGB, RF, SVM, NN, ENPLR	Electronic data records – patient characteristics, surgical variables	Patient health assessment after THA	Kunze et al. [22]
NN, SGB, SVM, RF, KNN, ENPLR	Patient-specific characteristics	quantifying the risk of revision THR	Klement et al. [23]
ML with operator (LASSO) model	Patient demographics, patient-reported health, cognitive appraisal processes and surgical approach	Prediction of post-operative THA	Sniderman et al. [24]
<i>Pre-operative decision-making</i>			
ANN, LR	Electronic data records – patient characteristics, surgical variables	Mortality rate prediction	Lin et al. [25]
SGB, RF, SVM, NN, ENPLR	Patient, laboratory characteristics, medications, etc	Predict opioid use after THA	Karhade et al. [26]
LR, RF, AdaBoost, gradient boosting machines, XGBoost	Electronic data records – patient characteristics, surgical variables	To determine perioperative risks	Shah et al. [31]
Multivariable LR, ANN, RF	Features including anesthesia type, body mass index, age, ethnicity, white blood cell count, sodium level, and alkaline phosphatase	Predicting same day discharge	Zhong et al. [32]
<i>Clinical decision-making</i>			
K-means clustering	Gait or sensor data and electronic data records	Predicting clinical outcomes or patient-reported outcome measures (PROM)	Bini et al. [33]
LASSO, RF, linear SVM	Hospital registry data	Predict 2-year postsurgical MCIDs	Fontana et al. [34]
SVM, LDA	Questionnaires or gait analysis	Fall risk assessment	Polus et al. [35]
Naïve Bayesian	Demographics and comorbidities	Prediction of length of stay and payment	Ramkumar et al. [36]
<i>THR monitoring</i>			
linear regression, CART, M5, random forest, gradient boosting, ANN, SVM, k-NN	Polyethylene wear data of prosthetic hip implants	Predict polyethylene wear rate	Borjali et al. [37]
<i>Simulation models</i>			
FE and Gaussian models	CT scans, cohort of individual subjects	Primary stability assessment subject to micromotions	Al-Dzirri et al. [40]



**Table 3** (continued)

ML algorithm	Dataset	Outcome	Reference
Simulation, boosted classification, regression trees (CART) models	Average treatment effects (ATEs) on EQ-5D-3L, 6 months after THR in male patients with hybrid versus cementless hip prostheses	Health-related quality of life (HRQOL) assessment	Kreif et al. [41]
Materialize MIMICS, RF, GB	Quantitative biomechanical and bone quality data from CT, electromyography, gait analysis	Pre- and post-surgery decision-making	Ricciardi et al. [43]
3DSlicer, CNN	CT scans and x-ray images	To reconstruct 3D images from 2D X-rays	Almeida et al. [44]
FE model, ANNs, SVMs	Hip implant geometric factors, 3D FE models	To optimize hip implant	Cilla et al. [45]



**Fig. 7** Summary of the factors influencing implant failure

stand out. Utilizing image-based models might not always be viable for routine monitoring because there is an operator dependency for data (image scan) collection and validation. To guarantee preventive care in this digital era, an automated implant monitoring tool that can aid in clinical diagnosis should be in place. For this to be effective, the model should be trained with a dataset that can be taken regularly during the patient’s daily activities. Data options can be acoustic emission or gait datasets which can reveal precise tribocorrosion modes to draw accurate conclusions on THA implant failure predictions.

### 3 An in vitro case study: Hip implant tribocorrosion prediction through ML models

The term bioacoustics in general represents the sounds produced by living organisms, whereas bioacoustic signals (BAS) imply the family of signals produced by the human body [56]. In medical diagnostics, BAS corresponds to the audible outcome; however, there are inaudible acoustic signals that can contribute to the early diagnosis of any failures of human implants. Acoustic emission (AE) signal is one of those signals that originated as stress waves due to any mechanical loading event acting on a body [7, 40].

In our previous studies [40, 52], we have reported the effectiveness of using AE to predict bio-tribocorrosion in dental and orthopedic implants. Considering the fact that AE is exclusively focused on any active defects, any tribocorrosion event on the implant interface can lead to the initiation of AE signals. These tribocorrosion events can lead to structural damage and permanent failure of the installed implant. Previous study results proved that tribocorrosion assessment

is possible with AE technique. The comparison between selected features indicates that higher absolute energy, higher friction coefficient, and lower corrosion potential are indications of increased corrosion [40]. Another computational model study confirmed the impact of tissue thickness in AE signal transmission and its detectability at the skin surface or AE receiver sensor [53].

This work is aimed to advance the previous results in [40] to confirm the efficacy of a machine learning (ML) algorithm to categorize the implant risk levels at the head-cup interface and accurately predict any tribocorrosion event. The significance of this work includes early detection of implant failure, which can help in determining preventive measures to avoid the chances of secondary total hip replacement (THR) revision surgery. This model is designed to predict the damage risk at the pin in the pin-on-ball setup, as shown in Fig. 8.

### 3.1 Methods

#### 3.1.1 Acoustic emission (AE)

Effectiveness of AE in the implant (hip, dental) tribocorrosion assessment is reported in our previous research [40, 52] and is the motive behind the thought of an automated structural monitoring tool. AE is a stress wave that is generated in the material due to any deformation or defects. When compared to ultrasound signals, AE exclusively listens for active defects or deformations and is sensitive to its subsequent defect activity [57]. An added advantage of AE signals is their non-invasive and non-intrusive nature, offering data capturing through an AE sensor placed on the skin/surface. It also opens up the possibility of deploying in vivo models

where the AE data includes different tribocorrosion elements such as wear, friction, and corrosion. AE signal feature analysis considering amplitude, frequency, energy, and count also reveals its usefulness in tribocorrosion assessment [7, 40, 52].

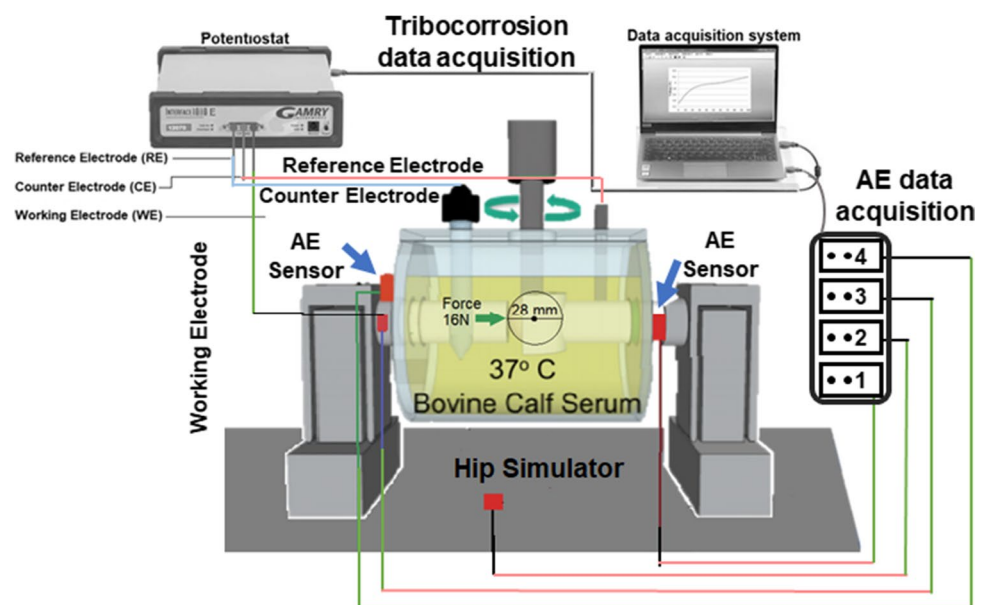
#### 3.1.2 Experimental approach

The experimental model includes a hip simulator under physiological conditions with a pin-on-ball setup to mimic the head-cup interface (Fig. 8). This model takes tribocorrosion and AE data from CoCrMo and Ti6Al4V pin (working electrode) versus saturated calomel electrode (reference electrode) for each experiment as input. Material selection for the working electrode was based on its use in THR with the hip simulator capable of maintaining physiological conditions closer to in vivo settings. Experimental data gathered include (a) tribocorrosion data: corrosion potential and friction coefficient, collected through the Gamry Echem Analyst tool and (b) acoustic emission data from the AE sensors and associated data acquisition system by the MISTRAS group, which facilitates easy handling of AE data. Our previous evaluation [40] considered corrosion potential, friction coefficient, weight loss, and AE absolute energy features as input to the preliminary ML model.

#### 3.1.3 Machine learning approach

Since the accuracy and performance of an ML model resides solely on the input data, proper attention should be given to its selection and processing. Tribocorrosion and AE analyses consider selected features such as corrosion potential, friction coefficient, weight loss, AE absolute energy in

**Fig. 8** Schematic diagram of the in vitro experimental setup, the tribocorrosion apparatus simulating pin-on-ball or head-cup interface model, which includes systems for mechanical, electrochemical, and acoustic emission data acquisition



determining the tribocorrosion failure. The collected data is preprocessed and given to the ML algorithm as an initial step. Selected features were then evaluated against the intensity of tribocorrosion occurrence (knowledge-based approach) and generated labels such as low, medium, and high for each occurrence. However, AE absolute energy data obtained are time-domain, which is further processed to identify the cumulative energy values sufficient to display the intensity of tribocorrosion. Once the data set is ready, it is sent to the scaling module that helps to normalize the independent feature sets.

As a result, 276 data points were prepared for each feature considering 92 data points per low, medium, and high corrosion groups. Following the general ML development approach, 80% were treated as training data, and the remaining 20% for testing purposes. Since implant tribocorrosion prediction using the ML approach is novel, identifying the efficacy of ML models that can function well with this dataset has great significance. Standardized ML algorithms such as LR, LDA, k-NN, DT, SVC, and RF models were selected for this study, and its accuracy in predicting tribocorrosion risk levels in the selected samples was confirmed. Other ML models, such as reinforcement learning or neural network algorithms, are still under consideration. ML section in this study contains 3 modules: (a) determining suitable ML algorithm, (b) feature importance study, (c) categorization or prediction of tribocorrosion occurrence. Based on the selected tasks, the results can be algorithm accuracy measured in percentage, feature importance measure, and prediction of implant failure level, respectively.

RF model, an ensemble of DTs, learns from random subsets of data, and each subtree delivers separate

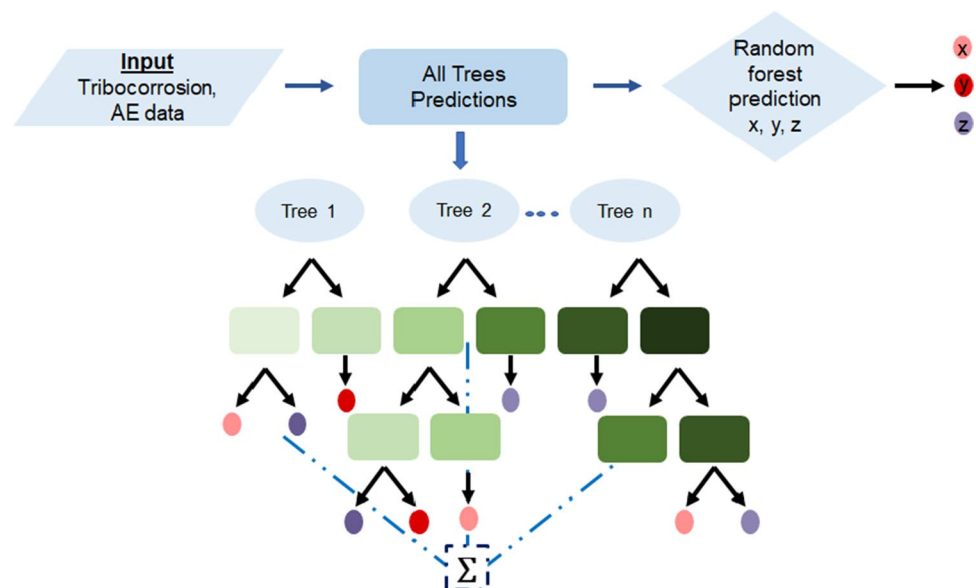
predictions or outcomes. The final decision from an RF model resides on the majority or average of the outcome from the subtrees (Fig. 9). That is, an RF model is an ensemble of a family of classifier DTs such as  $h(x|\Theta_1)$ ,  $h(x|\Theta_2)$ , ...  $h(x|\Theta_k)$  centered on classification and regression trees with parameters  $\Theta_k$  randomly chosen from the random model vector  $\Theta$ . Here,  $h_k(x)$  represents a single decision tree. The final classification ( $x$ ) is centered on the outcome of each tree where the class with the majority succeeds. The RF model is taken into account in the following modules of this study, considering its advantages over other selected algorithms, including the ability to handle multiple features without overfitting, decisions based on the majority outcome of multiple decision trees, the ability to deal with categorical and multiclass data, and supports for automatic feature interaction with minimal or no real-time execution [58–62].

### 3.2 Results

#### 3.2.1 Determining the suitability of ML algorithms

This step evaluates the accuracy of selected ML models on tribocorrosion prediction based on the four selected features. This module functions on a subset of data points from all the feature groups. Out of the algorithms that we screened, random forest (RF), decision tree (DT), and nearest-neighbor (k-NN) show promising results, as shown in Fig. 10. This step helped us to determine the RF model as the most suitable algorithm (from the selected ones) for predicting tribocorrosion occurrence.

**Fig. 9** A random forest model for 3 groups indicating x – low, y – medium, z – high corrosion class labels



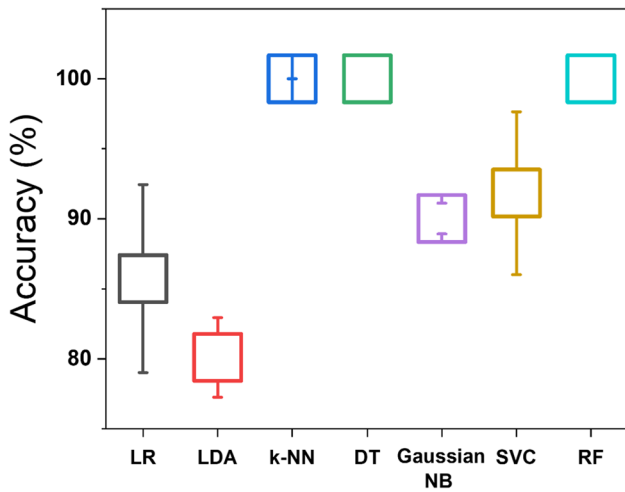


Fig. 10 Comparing classification accuracy for various ML algorithms for the tribocorrosion and AE data sets

### 3.2.2 Tribocorrosion feature importance study

Identifying contributing features for ML models is highly recommended, especially when dealing with multiple large sorts of features. The previous module confirms that the RF model is suitable for tribocorrosion prediction, while this section highlights the relevance of each feature in determining the tribocorrosion risk level for the ML model. In other words, utilizing the RF model’s feature importance matrix, this module determines which features contribute more and help decide the tribocorrosion class label. Figure 11 shows that AE absolute energy has almost similar relevance to other features in categorizing tribocorrosion risk levels for hip implant materials. This corroborates our previous study’s

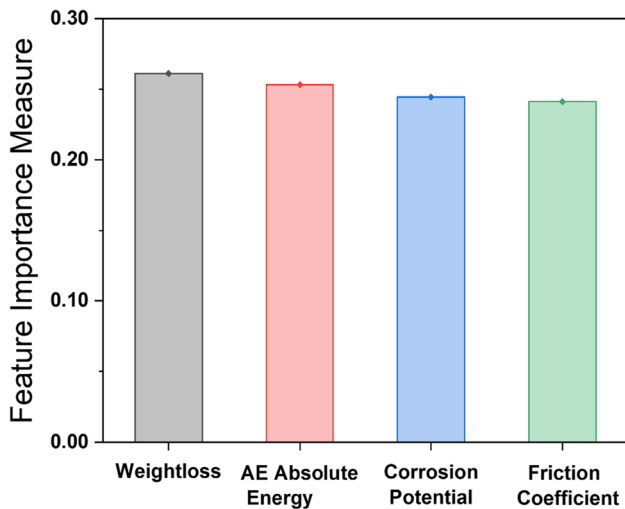


Fig. 11 Feature importance measure calculated using RF algorithm

finding that depicts AE’s usability in tribocorrosion prediction [40].

### 3.2.3 RF model: Determining the tribocorrosion risk level

Every ML model starts its learning through the training phase with pre-processed data. The learning phase starts with learning features and feature combinations for each class label or category. Likewise, the RF model is trained with a random choice of 80% of data from the entire dataset and learns from its training knowledge. Finally, the model performs classification and prediction of implant risk levels for the remaining 20% test dataset. Figure 12 is the circular grid representation of predictions over the actual corrosion labels. Actual corrosion labels indicate the class labels of the original test dataset, while prediction indicates the class labels predicted by the model for the test data. Here we can see that most of the predicted results in blue overlay the actual red color labels, which proves that most of the predicted tribocorrosion occurrences are accurately grouped under the desired category or class label.

Another way to interpret the ML model’s performance is through the confusion matrix (Fig. 13). Here the main diagonal cells (light-colored regions) indicate accurate classifications, while other cells show the incorrectly classified occurrences. Accuracy in percentage (%) is another important performance measure, and an accuracy of 94.81% indicates the developed ML model can derive accurate and effective decisions based on its training knowledge (Table 4).

RF model’s performance measures reveal the possibility of deploying ML models to determine implant tribocorrosion risk levels. Since it is a lab-scale in vitro study, the dataset is minimal in size from an ML model perspective.

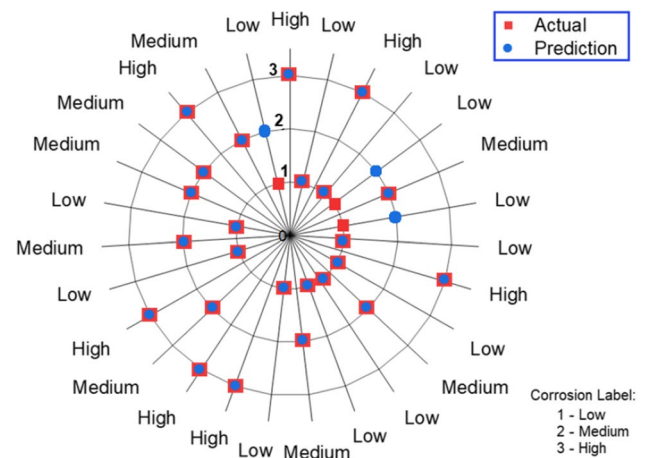
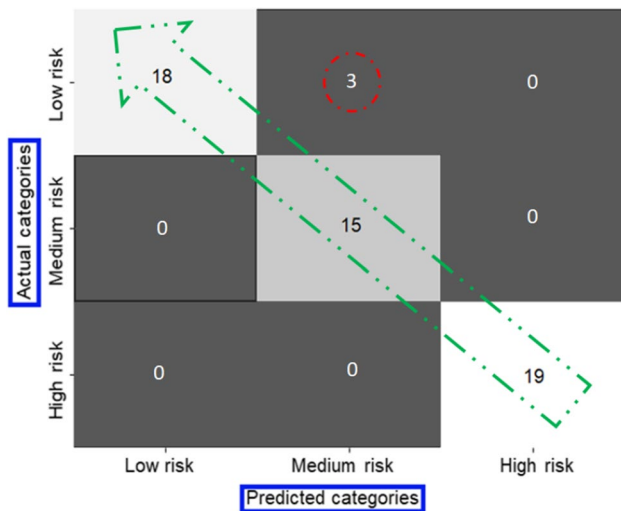


Fig. 12 Representative plot showing the RF prediction results of corrosion labels. Here three occurrences of low corrosion classes are incorrectly classified under the medium corrosion group



**Fig. 13** Confusion matrix or error matrix representing performance of the model

**Table 4** Model performance

Measure	Value
Mean absolute error	0.05
Mean squared error	0.02
Root mean squared error	0.14
Accuracy	94.81%

Even with this limitation, results guarantee the efficacy of ML models in tribocorrosion prediction.

### 4 Discussion

As bio-tribocorrosion is a major clinical concern in biomedical implants, so its monitoring in in vivo or clinical environments is highly significant. With the lack of a routine post-surgery implant monitoring facility, the lifetime of THA implants is ambiguous. A continuous monitoring technique should be adopted to efficiently interpret data during patients’ daily activities to deter such situations and facilitate the preventive maintenance feasibility of THA implants. Our previous studies [2–5] validated the efficacy of the AE method for tribocorrosion monitoring in an in vitro experimental setup and backed the necessity for such a continuous monitoring scheme. However, the present study goes beyond previous work by setting a higher bar by developing an AE-based continuous structural monitoring tool for THA implants utilizing the applicability of ML models. The novelty of the current work is the ML model itself, which accepts tribocorrosion features as input and predicts failure levels of THA implants.

Feature importance study confirms that AE features are best suited to determine tribocorrosion events in the headcup interface. However, continuous time-domain data, like the data collected during daily patient activities, should also be considered for training the model. This dataset can be AE signals indicating tribocorrosion behavior and gait analysis representing kinetics and kinematics of locomotion in biomedical implants. A similar study discussed by Dey et al. used biomechanical data for predicting the dynamics and ankle kinematics during level walking with SVR models [63].

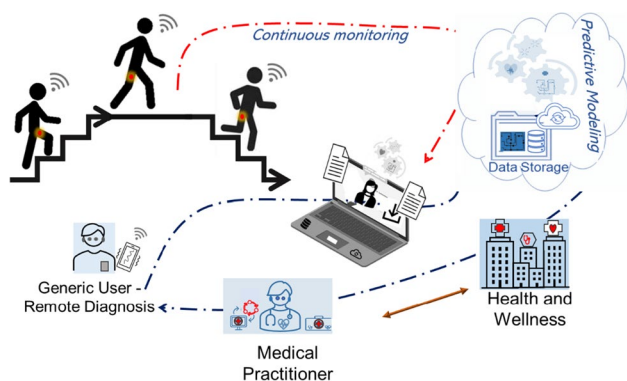
The proposed study supports the practicability of predictive models in tribocorrosion monitoring with a limited data set from an ML perspective. To establish this as a predictive modeling tool for bio-tribocorrosion detection and improve its accuracy, more research should be conducted to identify the efficacy of other continuous AE features and the best ML algorithm. As part of the prospect research, we are considering ideas around time-domain AE, biomechanical dataset, and tribocorrosion event localization. In the future, this, in turn, can contribute to developing an intelligent predictive modeling tool that can forecast and detect active damage mechanisms, especially due to tribocorrosion, the onset of failures, and the event or source location in THA implants.

With this ML model, we validated the relevance of AE features and tribocorrosion features, in predicting the tribocorrosion behavior of CoCrMo or Ti6Al4V samples. RF algorithm came out to be best suited for tribocorrosion datasets from the ML models selected for testing, where categorization and prediction of tribocorrosion damage levels have an accuracy greater than 90% [64].

Our recent simulation study confirmed the impact of tissue thickness on AE signal transmission and the feasibility of detecting AE signals at the skin surface [53]. Practical challenges such as AE signal transmission from the implant to the skin in patient trials; handling of large data set to feature the AE signal; the need for collaborative efforts between bioengineers, material scientists, computational experts, clinicians; and finally, the clinical trial of AE with patients with implant history could be constraints for prospect in-service application deployment. The key notion behind this study is to develop an automated home-based diagnosis tool (Fig. 14) that supports continuous hip implant monitoring in a point-of-care home-based setting, enabling data capture during daily patient activities and integrating data into an on-prem or cloud data store.

### 5 Conclusions

State-of-the-art AI-based ML (AI/ML) models can revolutionize digital orthopedics, especially with in vivo monitoring of implants.



**Fig. 14** Schematic representation of prospect point-of-care home-based continuous hip implant monitoring system

- Proposed AE feature-based ML model is an innovative method to continuously monitor active damage mechanisms associated with THA implants.
- Selection of appropriate AE signal features is key to this model's effectiveness, which guarantees an AI-based, non-invasive monitoring technique in real clinical applications.
- Understanding the practical challenges and effectively working around the constraints is crucial for developing and deploying continuous hip implant monitoring in a point-of-care home-based application.
- The model needs to be trained with a larger dataset of real-world patient data (acoustic emission and gait analysis) to facilitate higher efficiency and accurate outcomes.
- By considering appropriate algorithms and datasets, AI/ML models can revolutionize digital orthopedics, especially in in vivo monitoring of implants.

**Abbreviations** AI: Artificial intelligence; AE: Acoustic emission; ACS—NSQIP: American College of Surgeons—National Surgical Quality Improvement Program; AJRR: American Joint Replacement Registry; AP: Anteroposterior; ANN: Artificial neural network; BAS: Bioacoustic signals; CT: Computed tomography; CNN: Convolutional neural network; ENPLR: Elastic net penalized logistic regression; EHR: Electronic health records; HOOS: Hip disability and osteoarthritis outcome score; KNN: K-nearest neighbors; LASSO: Least Absolute Shrinkage Selection Operator; LDA: Linear discriminant analysis; LR: Logistic regression; ML: Machine learning; MRI: Magnetic resonance imaging; MCID: Minimally clinically important difference; NN: Neural network; PROM: Patient-reported outcome measures; PMCT: Post-mortem computed tomography; RF: Random forest; SGB: Stochastic gradient boosting; SVM: Support vector machine; THA: Total hip arthroplasty; THR: Total hip replacement; TJA: Total joint replacement surgery—arthroplasty; TKA: Total knee arthroplasty; US: Ultrasound

**Acknowledgements** The authors acknowledge Christine Lee, MS, 2020, Department of Biomedical Engineering UIC Chicago, Blazer Foundation for the Regenerative Medicine and Disability Research Lab

(RMDR) at the Department of Biomedical Sciences UIC College of Medicine at Rockford.

## Declarations

**Conflict of interest** The authors declare no competing interests.

## References

1. (2021) Digitalization in healthcare empowers patients. *Healthcare IT News*. <https://www.healthcareitnews.com/news/digitalization-healthcare-empowers-patients>. Accessed 25 Feb 2022
2. Gruson D, Helleputte T, Rousseau P, Gruson D (2019) Data science, artificial intelligence, and machine learning: opportunities for laboratory medicine and the value of positive regulation. *Clin Biochem* 69:1–7. <https://doi.org/10.1016/j.clinbiochem.2019.04.013>
3. Lee J-G, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N (2017) Deep learning in medical imaging: general overview. *Korean J Radiol* 18(4):570–584. <https://doi.org/10.3348/kjr.2017.18.4.570>
4. Nunamaker DM (2019) Orthopedic implant failure. In: *Equine fracture repair*. Wiley, pp 831–834. <https://doi.org/10.1002/9781119108757.ch46>
5. Delaunay C, Hamadouche M, Girard J, Duhamel A (2013) What are the causes for failures of primary hip arthroplasties in France? *Clin Orthop* 471(12):3863–3869. <https://doi.org/10.1007/s11999-013-2935-5>
6. Aqil, Shah N (2020) Diagnosis of the failed total hip replacement. *J Clin Orthop Trauma* 11(1):2–8. <https://doi.org/10.1016/j.jcot.2019.11.003>
7. Remya AR, Vishwash B, Lee C et al (2020) Hip implant performance prediction by acoustic emission techniques: a review. *Med Biol Eng Comput* 58:1637–1650. <https://doi.org/10.1007/s11517-020-02202-z>
8. Pei GX, Zhang YZ (2018) The establishment of digital orthopedics. In: Pei G (ed) *Digital orthopedics*. Springer Netherlands, Dordrecht, pp 3–7. [https://doi.org/10.1007/978-94-024-1076-1\\_1](https://doi.org/10.1007/978-94-024-1076-1_1)
9. Nam D, Barrack RL, Potter HG (2014) What are the advantages and disadvantages of imaging modalities to diagnose wear-related corrosion problems? *Clin Orthop* 472(12):3665–3673. <https://doi.org/10.1007/s11999-014-3579-9>
10. van den Heuvel L et al (2020) Quadruple decision making for Parkinson's disease patients: combining expert opinion, patient preferences, scientific evidence, and big data approaches to reach precision medicine. *J Park Dis* 10(1):223–231. <https://doi.org/10.3233/JPD-191712>
11. Kang Y-J, Yoo J-I, Cha Y-H, Park CH, Kim J-T (2020) Machine learning-based identification of hip arthroplasty designs. *J Orthop Transl* 21:13–17. <https://doi.org/10.1016/j.jot.2019.11.004>
12. Karnuta JM, Luu BC, Roth AL, Haerberle HS, Chen AF, Iorio R, Schaffer JL, Mont MA, Patterson BM, Krebs VE, Ramkumar PN (2021) Artificial intelligence to identify arthroplasty implants from radiographs of the hip. *J Arthroplasty* 36(7 Supplement):S290–S294.e1. <https://doi.org/10.1016/j.arth.2020.11.015>
13. Borjali A, Chen AF, Muratoglu OK, Morid MA, Varadarajan KM (2020) Detecting total hip replacement prosthesis design on plain radiographs using deep convolutional neural network. *J Orthop Res* 38(7):1465–1471. <https://doi.org/10.1002/jor.24617>
14. Murphy M, Killen C, Burnham R, Sarvari F, Wu K, Brown N (2021) Artificial intelligence accurately identifies total hip arthroplasty implants: a tool for revision surgery. *Hip Int J Clin*

- Exp Res Hip Pathol Ther: 1120700020987526. <https://doi.org/10.1177/1120700020987526>
15. Peña-Solórzano A, Albrecht DW, Bassed RB, Gillam J, Harris PC, Dimmock MR (2020) Semi-supervised labelling of the femur in a whole-body post-mortem CT database using deep learning. *Comput Biol Med* 122:103797. <https://doi.org/10.1016/j.combiomed.2020.103797>
  16. Shah RF, Bini SA, Martinez AM, Padoia V, Vail TP (2020) Incremental inputs improve the automated detection of implant loosening using machine-learning algorithms. *Bone Jt J*. Available: <https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.102B6.BJJ-2019-1577.R1>. Accessed 03 Mar 2022 [Online]
  17. Murphy EA, Killen C, Burnham R, Sarvari F, Wu K, Brown N (2022) Machine learning outperforms clinical experts in classification of hip fractures. *Sci Rep* 12(1):1. <https://doi.org/10.1038/s41598-022-06018-9>
  18. Jodeiri A, Zoroofi RA, Hiasa Y, Takao M, Sugano N, Yoshinobu S, Otake Y (2020) Fully automatic estimation of pelvic sagittal inclination from anterior-posterior radiography image using deep learning framework. *Comput Methods Programs Biomed* 184:105282. <https://doi.org/10.1016/j.cmpb.2019.105282>
  19. Milimonfared R, Oskoue RH, Taylor M, Solomon LB (2018) An intelligent system for image-based rating of corrosion severity at stem taper of retrieved hip replacement implants. *Med Eng Phys* 61:13–24. <https://doi.org/10.1016/j.medengphy.2018.08.002>
  20. Rouzrokh P, Ramazanian T, Wyles CC, Philbrick KA, Cai JC, Taunton MJ, Kremers HM, Lewallen DG, Erickson BJ (2021) Deep learning artificial intelligence model for assessment of hip dislocation risk following primary total hip arthroplasty from postoperative radiographs. *J Arthroplasty* 36(6):2197–2203.e3. <https://doi.org/10.1016/j.arth.2021.02.028>
  21. Rouzrokh P, Wyles CC, Philbrick KA, Ramazanian T, Weston AD, Cai JC, Taunton MJ, Lewallen DG, Berry DJ, Erickson BJ, Kremers HM (2021) A deep learning tool for automated radiographic measurement of acetabular component inclination and version after total hip arthroplasty. *J Arthroplasty* 36(7):2510–2517.e6. <https://doi.org/10.1016/j.arth.2021.02.026>
  22. Kunze KN, Karhade AV, Sadauskas AJ, Schwab JH, Levine BR (2020) Development of machine learning algorithms to predict clinically meaningful improvement for the patient-reported health state after total hip arthroplasty. *J Arthroplasty* 35(8):2119–2123. <https://doi.org/10.1016/j.arth.2020.03.019>
  23. Klemm, Laurencin S, Alpaugh K, Tirumala V, Barghi A, Yeo I, Subih MA, Kwon YM (2022) The utility of machine learning algorithms for the prediction of early revision surgery after primary total hip arthroplasty. *JAAOS - J Am Acad Orthop Surg*: <https://doi.org/10.5435/JAAOS>. <https://doi.org/10.5435/JAAOS-D-21-01039>
  24. Sniderman J, Stark RB, Schwartz CE, Imam H, Finkelstein JA, Nousiainen MT (2021) Patient factors that matter in predicting hip arthroplasty outcomes: a machine-learning approach. *J Arthroplasty* 36(6):2024–2032. <https://doi.org/10.1016/j.arth.2020.12.038>
  25. Lin C-C, Ou Y-K, Chen S-H, Liu Y-C, Lin J (2010) Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture. *Injury* 41(8):869–873. <https://doi.org/10.1016/j.injury.2010.04.023>
  26. Karhade AV, Schwab JH, Bedair HS (2019) Development of machine learning algorithms for prediction of sustained postoperative opioid prescriptions after total hip arthroplasty. *J Arthroplasty* 34(10):2272–2277.e1. <https://doi.org/10.1016/j.arth.2019.06.013>
  27. Harris HS, Kuo AC, Bozic KJ, Lau E, Bowe T, Gupta S, Giori NJ (2018) American joint replacement registry risk calculator does not predict 90-day mortality in veterans undergoing total joint replacement. *Clin Orthop* 476(9):1869–1875. <https://doi.org/10.1097/CORR.0000000000000377>
  28. Harris HS, Kuo AC, Bowe T, Gupta S, Nordin D, Giori NJ (2018) Prediction models for 30-day mortality and complications after total knee and hip arthroplasties for veteran health administration patients with osteoarthritis. *J Arthroplasty* 33(5):1539–1545. <https://doi.org/10.1016/j.arth.2017.12.003>
  29. Harris HS, Kuo AC, Weng Y, Trickey AW, Bowe T, Giori NJ (2019) Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? *Clin Orthop* 477(2):452–460. <https://doi.org/10.1097/CORR.0000000000000601>
  30. Edelstein I, Kwasny MJ, Suleiman LI, Khakhkhar RH, Moore MA, Beal MD, Manning DW (2015) Can the American College of Surgeons risk calculator predict 30-day complications after knee and hip arthroplasty? *J Arthroplasty* 30(9 Supplement):5–10. <https://doi.org/10.1016/j.arth.2015.01.057>
  31. Shah A, Devana SK, Lee C, Kianian R, van der Schaar M, SooHoo NF (2021) Development of a novel, potentially universal machine learning algorithm for prediction of complications after total hip arthroplasty. *J Arthroplasty* 36(5):1655–1662.e1. <https://doi.org/10.1016/j.arth.2020.12.040>
  32. Zhong H, Poeran J, Gu A, Wilson LA, Della Valle AG, Memtsoudis SG, Liu J (2021) Machine learning approaches in predicting ambulatory same day discharge patients after total hip arthroplasty. *Reg Anesth Pain Med* 46(9):779–783. <https://doi.org/10.1136/rapm-2021-102715>
  33. Bini SA, Shah RF, Bendich I, Patterson JT, Hwang KM, Zaid MB (2019) Machine learning algorithms can use wearable sensor data to accurately predict six-week patient-reported outcome scores following joint replacement in a prospective trial. *J Arthroplasty* 34(10):2242–2247. <https://doi.org/10.1016/j.arth.2019.07.024>
  34. Fontana MA, Lyman S, Sarker GK, Padgett DE, MacLean CH (2019) Can machine learning algorithms predict which patients will achieve minimally clinically important differences from total joint arthroplasty? *Clin Orthop* 477(6):1267–1279. <https://doi.org/10.1097/CORR.0000000000000687>
  35. Polus JS, Bloomfield RA, Vasarhelyi EM, Lanting BA, Teeter MG (2021) Machine learning predicts the fall risk of total hip arthroplasty patients based on wearable sensor instrumented performance tests. *J Arthroplasty* 36(2):573–578. <https://doi.org/10.1016/j.arth.2020.08.034>
  36. Ramkumar PN, Navarro SM, Haeberle HS, Karnuta JM, Mont MA, Iannotti JP, Patterson BM, Krebs VE (2019) Development and validation of a machine learning algorithm after primary total hip arthroplasty: applications to length of stay and payment models. *J Arthroplasty* 34(4):632–637. <https://doi.org/10.1016/j.arth.2018.12.030>
  37. Borjali A, Monson K, Raeymaekers B (2019) Predicting the polyethylene wear rate in pin-on-disc experiments in the context of prosthetic hip implants: Deriving a data-driven model using machine learning methods. *Tribol Int* 133:101–110. <https://doi.org/10.1016/j.triboint.2019.01.014>
  38. Gibon E, Amanatullah DF, Loi F, Pajarinen J, Nabeshima A, Yao Z, Hamadouche M, Goodman SB (2017) The biological response to orthopaedic implants for joint replacement: part I: metals. *J Biomed Mater Res B Appl Biomater* 105(7):2162–2173. <https://doi.org/10.1002/jbm.b.33734>
  39. Oladokun A, Pettersson M, Bryant M, Engqvist H, Persson C, Hall R, Neville A (2015) Fretting of CoCrMo and Ti6Al4V alloys in modular prostheses. *Tribol - Mater Surf Interfaces* 9(4):165–173. <https://doi.org/10.1179/1751584X15Y.0000000014>
  40. Lee, Zhang L, Morris D, Cheng KY, Ramachandran RA, Barba M, Bijukumar D, Ozevin D, Mathew MT (2021) Non-invasive early detection of failure modes in total hip replacements (THR) via acoustic emission (AE). *J Mech Behav Biomed Mater* 118:104484

41. (n.d.) What Is FEA | Finite element analysis? Documentation. SimScale. <https://www.simscale.com/docs/simwiki/fea-finite-element-analysis/what-is-fea-finite-element-analysis/>. Accessed 22 Jun 2022
42. Al-Dirini RMA, Martelli S, Taylor M (2020) Computational efficient method for assessing the influence of surgical variability on primary stability of a contemporary femoral stem in a cohort of subjects. *Biomech Model Mechanobiol* 19(4):1283–1295. <https://doi.org/10.1007/s10237-019-01235-0>
43. Kreif N, Gruber S, Radice R, Grieve R, Sekhon JS (2016) Evaluating treatment effectiveness under model misspecification: a comparison of targeted maximum likelihood estimation with bias-corrected matching. *Stat Methods Med Res* 25(5):2315–2336. <https://doi.org/10.1177/0962280214521341>
44. Viceconti M, Testi D, Gori R, Zannoni C, Cappello A, Nulldelollis A (2001) HIDE: a new hybrid environment for the design of custom-made hip prosthesis. *Comput Methods Programs Biomed* 64(2):137–144. [https://doi.org/10.1016/s0169-2607\(00\)00097-3](https://doi.org/10.1016/s0169-2607(00)00097-3)
45. Ricciardi, Jónsson H, Jacob D, Improta G, Recenti M, Gíslason MK, Cesarelli G, Esposito L, Minutolo V, Bifulco P, Gargiulo P (2020) Improving prosthetic selection and predicting BMD from biometric measurements in patients receiving total hip arthroplasty. *Diagn Basel Switz* 10(10):E815. <https://doi.org/10.3390/diagnostics10100815>
46. Almeida F, Astudillo P, Vandermeulen D (2021) Three-dimensional image volumes from two-dimensional digitally reconstructed radiographs: a deep learning approach in lower limb CT scans. *Med Phys* 48(5):2448–2457. <https://doi.org/10.1002/mp.14835>
47. Cilla M, Borgiani E, Martínez J, Duda GN, Checa S (2017) Machine learning techniques for the optimization of joint replacements: application to a short-stem hip implant. *PLoS One* 12(9):e0183755. <https://doi.org/10.1371/journal.pone.0183755>
48. Ramachandran RA, Chi S-W, Ozevin D, Barba M, Mathew MT (n.d.) Acoustic emission to monitor hip implant well-function: a computational approach to overcome the impact of surrounding tissue thickness, p. 1. ORS 2021 Annual Meeting. <https://www.ors.org/abstract-search/>. Accessed 5 Mar 2021
49. Wang S, Summers RM (2012) Machine learning and radiology. *Med Image Anal* 16(5):933–951. <https://doi.org/10.1016/j.media.2012.02.005>
50. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL (2018) Artificial intelligence in radiology. *Nat Rev Cancer* 18(8):500–510. <https://doi.org/10.1038/s41568-018-0016-5>
51. Clohisy JC, Calvert G, Tull F, McDonald D, Maloney WJ (2004) Reasons for revision hip surgery: a retrospective review. *Clin Orthop Relat Res* 429:188–192. <https://doi.org/10.1097/01.blo.0000150126.73024.42>
52. Barão VAR, Ramachandran RA, Matos AO, Badhe RV, Grandini CR, Sukotjo C, Ozevin D, Mathew M (2021) Prediction of tribocorrosion processes in titanium-based dental implants using acoustic emission technique: Initial outcome. *Mater Sci Eng C* 123:112000. <https://doi.org/10.1016/j.msec.2021.112000>
53. Ampadi Ramachandran R, Lee C, Zhang L et al (2022) Total hip replacement monitoring: numerical models for the acoustic emission technique. *Med Biol Eng Comput* 60:1497–1510. <https://doi.org/10.1007/s11517-022-02548-6>
54. Callahan A, Fries JA, Ré C, Huddleston JI, Giori NJ, Delp S, Shah NH (2019) Medical device surveillance with electronic health records. *NPJ Digit Med* 2:94. <https://doi.org/10.1038/s41746-019-0168-z>
55. Nich A, Behr J, Crenn V, Normand N, Mouchère H, d'Assignies G (2022) Applications of artificial intelligence and machine learning for the hip and knee surgeon: current state and implications for the future. *Int Orthop*. <https://doi.org/10.1007/s00264-022-05346-9>
56. Hadjileontiadis LJ, Rekanos IT, Panas SM (2006) Bioacoustic signals. In: Akay M (ed) *Wiley Encyclopedia of Biomedical Engineering*. Wiley, Hoboken, p ebs0109. <https://doi.org/10.1002/9780471740360.ebs0109>
57. Breyse D (ed) (2012) Non-destructive assessment of concrete structures: Reliability and limits of single and combined techniques: State-of-the-art report of the RILEM technical committee 207-INR, vol. 1. Springer Science & Business Media
58. (n.d.) Random Forests | SpringerLink. <https://link.springer.com/article/10.1023/a:1010933404324>. Accessed 01 Nov 2021
59. Kon M (n.d.) Mathematics of random forests. Available: <http://math.bu.edu/people/mkon/MA751/L19RandomForestMath.pdf>. Accessed 25 Mar 2022 [Online]
60. Varghese (2019) Comparative study on classic machine learning algorithms. Medium. <https://towardsdatascience.com/comparative-study-on-classic-machine-learning-algorithms-24f9ff6ab222>. Accessed 25 Mar 2022
61. Varghese D (2018) Comparative study on classic machine learning algorithms , Part-2. Medium. <https://medium.com/@dannymvarghese/comparative-study-on-classic-machine-learning-algorithms-part-2-5ab58b683ec0>. Accessed 25 Mar 2022
62. Kumar A (2020) Hold-out method for training machine learning models. *Data Analytics*. <https://vitalflux.com/hold-out-method-for-training-machine-learning-model/>. Accessed 02 Nov 2021
63. Dey S, Eslamy M, Yoshida T, Ernst M, Schmalz T, Schilling A (2019) A support vector regression approach for continuous prediction of ankle angle and moment during walking: an implication for developing a control strategy for active ankle prostheses. *IEEE Int Conf Rehabil Robot Proc* 2019:727–733. <https://doi.org/10.1109/ICORR.2019.8779445>
64. Ampadi Ramachandran R, Chi S-W, Ozevin D, Foucher K, Mathew MT (n.d.) Categorizing and predicting Tribocorrosion damage risks in THR implants deploying machine learning models: a preliminary approach. ORS 2022 Annual Meeting Paper No. 11, p 1. <https://www.ors.org/abstract-search/>. Accessed 05 Mar 2021

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



**Remya Ampadi Ramachandran** Ph.D. in Acoustics and speaker recognition, Postdoc researcher at the Department of Biomedical Engineering, University of Illinois Chicago. The research focused on biomedical implant monitoring through acoustic emission, signal processing, computational methods for biomedical applications, machine learning, and deep learning models for biomedical applications, Biotribocorrosion.



**Sheng-Wei Chi** Ph.D. in Civil and Environmental Engineering, Associate Professor at Department of Civil, Materials, and Environmental Engineering, University of Illinois Chicago. Research interests include Image-based computational methods for biomedical applications, Multiscale material modeling, Advanced numerical methods, including meshfree methods, collocation methods, and generalized finite element methods, large deformation and contact mechanics, fragment and impact simulations.



**P. Srinivasa Pai** Ph.D. in Mechanical Engineering, Professor at Department of Mechanical Engineering, NMAM Institute of Technology, Nitte, India. Research interests are tool condition monitoring acoustic emission, vibration monitoring, artificial neural networks maintenance engineering and management, wavelet and its applications to machining related areas, high-speed machining-related areas including wear studies, surface roughness analysis.



**Kharma Foucher** Ph.D. in Bioengineering, Associate Professor at the Department of Kinesiology and Nutrition, University of Illinois Chicago. The research focused on Biomechanics, motion analysis, musculoskeletal modeling of people with and without hip implants, and clinical outcomes.



**Didem Ozevin** Ph.D. in Materials Engineering, Associate Professor at Department of Civil, Materials, and Environmental Engineering, University of Illinois Chicago. Research interest focuses on sensor development to combine a variety of sensing elements and to develop new structural evaluation methods for understanding the damage states using a variety of structural health-monitoring techniques, such as acoustic emission and low-frequency ultrasounds. Acoustic emission and numerical models for biomedical applications.



**Mathew T. Mathew** Ph.D. in Mechanical Engineering. Associate Professor at Department of Biomedical Science, Director of Faculty Research, UIC School of Medicine, Rockford, IL, USA; Associate Professor, Department of Biomedical Engineering, University of Illinois Chicago. Leading the Regenerative Medicine and Disability lab at UIC Rockford. The research focused on Biomaterials, Biomedical implants, Corrosion, Tribocorrosion, Biotribocorrosion.