## Contents

### MINIREVIEWS

620  Coronavirus: An overview with special emphasis on COVID-19 outbreak with musculoskeletal manifestations  

629  Hip prosthetic loosening: A very personal review  
*Mjöberg B*

640  Jones type fifth metatarsal fracture fixation in athletes: A review and current concept  
*Albloushi M, Alshanqiti A, Qasem M, Abitbol A, Gregory T*

651  Effects of sclerostin antibody on bone healing  

660  Current concepts in the management of bisphosphonate associated atypical femoral fractures  

672  Virtual orthopedic assessment: Main principles and specific regions  
*Emara KM, Diab RA, Amr M, Mahmoud SA, Nour M, Abdelsalam A, Elshobaky MA, Farhan M, Gmeah M, Emara AK*

685  Machine learning in orthopaedic surgery  
*Lalehzarian SP, Gowd AK, Liu JN*

### ORIGINAL ARTICLE

#### Clinical and Translational Research

700  Trends and risk factors for opioid administration for non-emergent lower back pain  
*Gwam CU, Emara AK, Chughtai N, Javed S, Luo TD, Wang KY, Chughtai M, O’Gara T, Plate JF*

#### Retrospective Study

710  Correlation of stress radiographs to injuries associated with lateral ankle instability  
*Sy JW, Lopez AJ, Lausé GE, Deal JB, Lustik MB, Ryan PM*

### CASE REPORT

720  Atypical osteochondroma of the lumbar spine associated with suprasellar pineal germinoma: A case report  
*Suwak P, Barnett SA, Song BM, Heffernan MJ*
World Journal of Orthopedics

Monthly Volume 12 Number 9 September 18, 2021

ABOUT COVER

Editorial Board Member of World Journal of Orthopedics, Vasileios S Nikolaou, MD, MSc, PhD, Associate Professor, Surgeon, The Second Orthopaedic Department, School of Medicine, National and Kapodistrian University of Athens, Athens 14233, Greece. vassilios.nikolaou@gmail.com

AIMS AND SCOPE

The primary aim of World Journal of Orthopedics (WJO, World J Orthop) is to provide scholars and readers from various fields of orthopedics with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJO mainly publishes articles reporting research results and findings obtained in the field of orthopedics and covering a wide range of topics including arthroscopy, bone trauma, bone tumors, hand and foot surgery, joint surgery, orthopedic trauma, osteoarthropathy, osteoporosis, pediatric orthopedics, spinal diseases, spine surgery, and sports medicine.

INDEXING/ABSTRACTING

The WJO is now abstracted and indexed in PubMed, PubMed Central, Emerging Sources Citation Index (Web of Science), Scopus, China National Knowledge Infrastructure (CNKI), China Science and Technology Journal Database (CSTJ), and Superstar Journals Database. The 2021 edition of Journal Citation Reports® cites the 2020 Journal Citation Indicator (JCI) for WJO as 0.66. The WJO’s CiteScore for 2020 is 3.2 and Scopus CiteScore rank 2020: Orthopedics and Sports Medicine is 87/262.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: Lin-YuTong Wang; Production Department Director: Xiang Li; Editorial Office Director: Jin-Lei Wang.

NAME OF JOURNAL

World Journal of Orthopedics

ISSN

ISSN 2218-5836 (online)

LAUNCH DATE

November 18, 2010

FREQUENCY

Monthly

EDITORS-IN-CHIEF

Massimiliano Leigheb

EDITORIAL BOARD MEMBERS

http://www.wjgnet.com/2218-5836/editorialboard.htm

PUBLICATION DATE

September 18, 2021

COPYRIGHT

© 2021 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS

https://www.wjgnet.com/bpg/gerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS

https://www.wjgnet.com/bpg/gerinfo/288

PUBLICATION MISCONDUCT

https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE

https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS

https://www.wjgnet.com/bpg/gerinfo/239

ONLINE SUBMISSION

https://www.f6publishing.com

© 2021 Baishideng Publishing Group Inc. All rights reserved. 7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
E-mail: bpgoffice@wjgnet.com https://www.wjgnet.com
Machine learning in orthopaedic surgery

Simon P Lalehzarian, Anirudh K Gowd, Joseph N Liu

ORCID number: Simon P Lalehzarian 0000-0003-3406-3651; Anirudh K Gowd 0000-0001-7151-6459; Joseph N Liu 0000-0002-3801-8885.

Author contributions: All authors made significant contributions toward the preparation of this manuscript; Lalehzarian SP wrote the article, critically revised the article, and participated in the final approval of the version to be published; Gowd AK critically revised the article and participated in the final approval of the version to be published; Liu JN designed the work, critically revised the article, and was responsible for final approval of the version to be published.

Conflict-of-interest statement: All authors have no conflicts of interest to report.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/License

Simon P Lalehzarian, The Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL 60064, United States

Anirudh K Gowd, Department of Orthopaedic Surgery, Wake Forest Baptist Medical Center, Winston-Salem, NC 27157, United States

Joseph N Liu, USC Epstein Family Center for Sports Medicine, Keck Medicine of USC, Los Angeles, CA 90033, United States

Corresponding author: Joseph N Liu, MD, Assistant Professor, USC Epstein Family Center for Sports Medicine, Keck Medicine of USC, 1520 San Pablo St #2000, Los Angeles, CA 90033, United States. joseph.liu@med.usc.edu

Abstract

Artificial intelligence and machine learning in orthopaedic surgery has gained mass interest over the last decade or so. In prior studies, researchers have demonstrated that machine learning in orthopaedics can be used for different applications such as fracture detection, bone tumor diagnosis, detecting hip implant mechanical loosening, and grading osteoarthritis. As time goes on, the utility of artificial intelligence and machine learning algorithms, such as deep learning, continues to grow and expand in orthopaedic surgery. The purpose of this review is to provide an understanding of the concepts of machine learning and a background of current and future orthopaedic applications of machine learning in risk assessment, outcomes assessment, imaging, and basic science fields. In most cases, machine learning has proven to be just as effective, if not more effective, than prior methods such as logistic regression in assessment and prediction. With the help of deep learning algorithms, such as artificial neural networks and convolutional neural networks, artificial intelligence in orthopaedics has been able to improve diagnostic accuracy and speed, flag the most critical and urgent patients for immediate attention, reduce the amount of human error, reduce the strain on medical professionals, and improve care. Because machine learning has shown diagnostic and prognostic uses in orthopaedic surgery, physicians should continue to research these techniques and be trained to use these methods effectively in order to improve orthopaedic treatment.

Key Words: Artificial intelligence; Machine learning; Supervised learning; Unsupervised learning; Deep learning; Orthopaedic surgery

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.
INTRODUCTION

The application of artificial intelligence (AI) has taken our world by storm. AI has been used in many aspects of modern life such as recommendation systems used by Netflix, YouTube, and Spotify, search engines like Google, and social-media feeds like Facebook and Twitter[1]. Additionally, AI has entered the realm of medicine. For example, there is substantial evidence that AI performs on par or better than humans in various tasks such as analyzing medical images as well as correlating symptoms and biomarkers from electronic medical records with the characterization and prognosis of disease[2]. Specifically in orthopaedic surgery, certain subfields of AI have been successfully implemented to improve clinical decision making and patient care[3].

To better understand the utility of AI in orthopaedic surgery, some terms must first be defined. AI started as a theory that computers could eventually learn to perform tasks through pattern recognition with minimal to no human involvement[1,4,5]. Today, the definition has been adapted to include the application of algorithms that provide machines the ability to solve problems that traditionally required human intelligence [1,5]. AI, which is often used as an umbrella term, encompasses subfields such as machine learning (ML), which is defined as a series of mathematical algorithms that enable the machine to “learn” the relationship between the input and output data without being explicitly told how to do so[3]. Furthermore, machine learning contains the subfield of deep learning (DL) which can be used to find correlations without labelling, that are too complex to render using previous machine learning algorithms, by processing input data through artificial neural networks[6,7].

In the field of orthopaedic surgery, ML has been used for different applications such as fracture detection, bone tumor diagnosis, detecting hip implant mechanical loosening, and grading osteoarthritis (OA)[3]. As time goes on, the utility of AI and ML in orthopaedic surgery continues to grow and expand. The purpose of this review is to provide an understanding of the concepts of ML and a background of current and future orthopaedic applications of ML in risk assessment, outcomes assessment, imaging, and basic science fields.

WHAT IS ML?

ML focuses on developing automated computer systems that predict outputs through algorithms and mathematics[8,9]. Classic or conventional ML algorithms that are meant to extract knowledge from more tabulated data sets include decision trees, random forests, nearest neighbors, linear regression, support vector machine (SVM), and k-means clustering[3,10]. On the other hand, more recently developed DL algorithms and artificial neural networks (ANN) are used to extract knowledge from imaging data sets. Regardless of which algorithms are used, ML requires software to “learn” patterns or relationships from sets of empirical data. This “learning” can be achieved through three different means: supervised learning, unsupervised learning,
and reinforcement learning[11-13].

**Types of ML**

Supervised learning, also termed inductive learning, is the most prevalent type of ML and occurs when data is labeled to tell the machine exactly what patterns it should look for[14]. For example, if an ML algorithm is used to detect arthritis on a knee radiograph, the arthritic features must be manually identified and labeled by a human along with the label of whether the radiograph is an example of an arthritic or normal knee[1]. On the other hand, unsupervised learning, also termed deductive or analytic learning, occurs when the data is not labeled and the machine looks for patterns[14] (Figure 1). In continuing with the last example, the arthritic features in unsupervised learning would not be labeled and therefore the algorithm relies on self-organization[1]. Lastly, reinforcement learning acts more like a reward or punishment system. Unlike supervised learning which makes data available at the beginning of the task, reinforcement learning uses feedback about the correctness after the task has been completed[15]. Usually, supervised learning is used because it requires the least amount of data and thus the least amount of time to learn.

**DL and ANN**

DL is modeled after the human brain’s neural connections via complex and layered algorithms termed ANN[1,14]. The complex layering allows the algorithm to learn more complex and subtle patterns compared to more simple one or two layer networks[5,16,17]. Two known models of deep learning within the ANN include convolutional neural network (CNN) and recurrent neural network. The main type, CNN, has two main functions: (1) to extract features from imaging; and (2) classification[3]. The CNN extraction feature relies on the idea that filters learned on a small subset of a larger image to detect certain features can also be applied to other parts of the larger image in order to detect the same feature at different locations[3]. The CNN starts by searching for simple features in an image and then pools these simple features together to extract more complicated high-level features[3,8] (Figure 2). For the classification feature, the CNN acts as a classic neural network that combines all the high-level feature maps (generated from the aforementioned filters) from the deepest convolution layer and uses them to output a classification score[3]. During training, the CNN is presented with a series of images that have known classifications; the CNN must make a classification decision for each image and then calculate the classification error by comparing its classification decision with the known classification of the image. Through this training process, the CNN is able to update its learnable parameters and make classification decisions on images never before seen[3].

DL algorithms have been successfully applied to complex problems to improve diagnostic accuracy and speed, flag the most critical and urgent patients for immediate attention, reduce the amount of human error, reduce the strain on medical professionals, and improve orthopaedic care[3]. Specifically in orthopaedic surgery, the greatest application of DL is in image classification.

**RISK ASSESSMENT**

While ML has traditionally been used in medicine for rule-based approaches such as safe drug prescription, recent use of ML and DL in orthopaedic surgery has focused on clinical decision support such as risk assessment[14,19]. Currently, logistic regression is one of the most commonly used methods for identifying risk factors predictive of developing complications; however, in comparison, ANN allows for the identification of nonlinear patterns that make predictions more accurate[20-22].

Throughout orthopaedic literature, the application of ML and DL in risk assessment for various complications has been studied extensively (Table 1). For example, in Kim et al[23], ML models were used to predict mortality, venous thromboembolism, cardiac complications, and wound complications following posterior lumbar fusion. The ML models outperformed the American Society of Anesthesiologists (ASA) scores proving that ML can be more effective at predicting complications. Similarly, in Harris et al[24], ML was used to predict 30 day mortality and morbidity after total joint arthroplasty. While the ML model was found to be more accurate than standard models for cardiac complications and mortality, it was less effective for rarer complications such as re-operation and deep infection. More recently, Gowd et al[25] used supervised ML models to predict postoperative outcomes following total shoulder arthroplasty. ML algorithms outperformed the standard model for predicting adverse events,
Table 1 Summary of machine learning for orthopaedic surgery risk assessment

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bevevino et al [26]</td>
<td>ANN capable of accurately estimating the likelihood of amputation</td>
</tr>
<tr>
<td>Gowd et al [25]</td>
<td>Supervised ML outperformed ASA classification models in predicting adverse events, transfusion, extended length of stay, surgical site infection, return to operating room, and readmission</td>
</tr>
<tr>
<td>Harris et al [24]</td>
<td>ML was moderately accurate in predicting 30-d mortality and cardiac complications after elective primary TJA</td>
</tr>
<tr>
<td>Kim et al [23]</td>
<td>ANN more accurate than ASA in predicting mortality, VTE, cardiac and wound complications following posterior lumbar spine fusion</td>
</tr>
</tbody>
</table>


Figure 2 Schematic diagram of a basic convolutional neural network architecture [18]. Reused with permission. Citation: Phung VH, Rhee EJ. A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. App Sci 2019; 9: 4500.

transfusion, extended length of stay, surgical site infection, return to the operating room, and readmission. Furthermore, in risk assessment related to orthopaedic trauma, Bevevino et al [26] used a DL model to predict the likelihood of amputation based on 155 combat-related open calcaneal fractures and compared it to a standard logistical regression model. Twenty-six features with a proven or theoretical association with successful or unsuccessful limb salvage were analyzed; some of the features included were various patient demographics, mechanism of injury, wound size and location, and fracture type. Once again, the DL method was 30% more accurate and better suited to clinical use than the standard logistical regression model.

The orthopaedic literature shows that ML continuously outperforms more traditional legacy risk-stratification measures such as ASA classification, Charlson Comorbidity Index, and modified 5-item frailty index, in predicting complications
following a variety of orthopaedic procedures as well as identifying safe candidates for specific orthopaedic procedures like anterior cervical fusion and disectomy[25,27-29]. In all of the studies outlined above, the ML and DL algorithms outperformed the standard models indicating a higher level of accuracy for risk assessment[23-26]. Furthermore, comorbidity indices have previously been used to gauge perioperative risk, evaluate the need for postoperative admission, and determine prophylactic treatments[30-32]. With continued validation, ML algorithms may replace this paradigm. Whereas logistical regression models have typically been used for many years to predict risk, survival, mortality, and morbidity, the application of ML and DL, particularly in predicting the risk of complications following spine surgeries, joint surgeries, and orthopaedic trauma, shows much promise[14].

OUTCOMES ASSESSMENT

Over the last twenty years or so, the ability to predict outcomes has positively impacted medicine and patient care. From risk scores that guide anticoagulation (CHA2DS2) to the use of cholesterol medications (ASCVD), data-driven clinical predictions have become routine in medical practice[33]. Because ML has the ability to analyze large data sets, the accuracy of prediction significantly improves[33]. Specifically in orthopaedic surgery, recent literature has shown the utility of ML algorithms in outcomes assessment for orthopaedic oncology survival, patient-reported outcome measures (PROMs), hospital length of stay, and cost (Table 2).

When compared to ML algorithms, current prognostication models for survival following metastatic spinal disease are not built for estimation of short-term survival (30 to 90 d) and some studies even suggest a lack of accuracy in classic models[34-37]. In Janssen et al[38], authors compared a boosting ML algorithm to a classic scoring system and nomogram at 30 d, 90 d, and 1 year to study survival estimates in patients with long bone metastases. In all training data sets, the boosting ML algorithm was found to be far superior at each time point[38]. Paulino Pereira et al[37] conducted a similar study where they compared the boosting ML algorithm, nomogram, and the classic scoring system to predict survival in metastatic spine disease. In this study, the boosting ML algorithm was comparable to nomogram in its predictive ability for testing data sets. Thio et al[39] and Bongers et al[40] used ML methods on patient demographics, tumor characteristics, treatment, and outcome data to create an ML algorithm that could predict 5 year survival rate in patients with chondrosarcoma[39,40]. In the latter study, the authors found that the ML algorithm overestimated the survival rate in their data set, but when applied to a smaller data set, it overestimated survival to a lesser extent.

Within the last ten years, the concept of PROMs has gained rapid support in orthopaedic surgery as a way to measure healthcare quality and value[41]. The minimally clinically important difference (MCID) or the minimum change in PROM scores that patients perceive as clinically meaningful offers a threshold of score that portends clinical relevance[41-43]. Using predictive models to identify patients at risk of not achieving MCID is important for resource allocation as well as better monitoring especially for presurgical decision support[41]. Fontana et al[41] used three ML models to predict which patients would not achieve a MCID in four PROMs two years following total joint arthroplasty (TJA). When applied to presurgical registry data, the three ML models predicted 2-year postsurgical MCIDs with fair-to-good ability showing that ML has good predictive power in MCID following TJA. In another study, Menendez et al[44] used ML to understand sentiment by exploring the content of negative patient-experience comments after total shoulder arthroplasty (TSA). Through a ML based approach, they found that patient satisfaction was highly correlated to hospital environment, nontechnical skills, and delays. Menendez et al[44] showed the potential utility of AI and ML models to analyze post-surgical PROM surveys to determine quality and satisfaction after TSA.

A newer trend in orthopaedic surgery is using ML concepts to predict hospital length of stay as well as cost[45-47]. Today, ML models can be used to predict how long or how much a patient’s surgery will cost prior to the elective procedure[48,49]. Ramkumar et al[45,47] and Navarro et al[46] used ML techniques on preoperative big data to predict length of stay and patient-specific payments following total hip arthroplasty (THA) and total knee arthroplasty (TKA), respectively[45-47]. In both studies, the ML techniques showed excellent predictability in length of stay. As complexity of the case increased, accuracy for predicting payment decreased proportionately in THA. On the other hand, as complexity of the case increased in TKA, accuracy for
ML in orthopaedic surgery

Table 2 Summary of machine learning for orthopaedic surgery outcomes assessment

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bongers <em>et al.</em>[40]</td>
<td>ML algorithm overestimated ability to predict 5-year survival in patients with chondrosarcoma</td>
</tr>
<tr>
<td>Fontana <em>et al.</em>[41]</td>
<td>Used ML to demonstrate fair-to-good ability in predicting 2-year postsurgical MCID following TJA</td>
</tr>
<tr>
<td>Greenstein <em>et al.</em>[51]</td>
<td>Used EMR-integrated ANN to predict discharge disposition after TJA on small data set</td>
</tr>
<tr>
<td>Janssen <em>et al.</em>[38]</td>
<td>Boosting ML algorithm far superior in training data sets to classic scoring system and nomogram in predicting survival in patients with long bone metastases at 30 days, 90 days, and 1 year</td>
</tr>
<tr>
<td>Karnuta <em>et al.</em>[30]</td>
<td>Bayes ML algorithm demonstrated excellent accuracy in prediction of length of stay and cost of an episode of care for hip fracture</td>
</tr>
<tr>
<td>Menendez <em>et al.</em>[44]</td>
<td>Used ML on patient-narrative analysis to show patient satisfaction after TSA is linked to hospital environment, non-technical skills, and delays</td>
</tr>
<tr>
<td>Navarro <em>et al.</em>[46]</td>
<td>Created a valid ML algorithm that predicted length of stay and costs before primary TKA</td>
</tr>
<tr>
<td>Pereira <em>et al.</em>[55]</td>
<td>Boosting ML algorithm comparable to nomogram in its ability to predict survival in metastatic spine disease with testing data sets</td>
</tr>
<tr>
<td>Ramkumar <em>et al.</em>[45]</td>
<td>Created a valid and reliable ML algorithm that predicted length of stay and payment prior to primary THA</td>
</tr>
<tr>
<td>Ramkumar <em>et al.</em>[47]</td>
<td>Developed several ML based models for primary LEA that preoperatively predict cost, length of stay, and discharge disposition</td>
</tr>
<tr>
<td>Thio <em>et al.</em>[39]</td>
<td>Created a high performing ML algorithm that could predict 5-year survival in patients with chondrosarcoma</td>
</tr>
</tbody>
</table>

ML: Machine learning; MCID: Minimally clinically important difference; TJA: Total joint arthroplasty; EMR: Electronic medical record; ANN: Artificial neural network; TSA: Total shoulder arthroplasty; TKA: Total knee arthroplasty; THA: Total hip arthroplasty; LEA: Lower extremity arthroplasty.

Predicting costs increased by 3%, 10%, and 15% for moderate, severe, and extreme risk populations[46]. Similarly, Karnuta *et al.*[30] used an ML algorithm on preoperative patient data to predict length of stay and cost after hip fracture; they found their ML algorithm to be 76.5% accurate for predicting length of stay and 79% accurate for predicting cost. Furthermore, Greenstein *et al.*[51] used ML to preoperatively predict the likelihood a patient will be discharged to a skilled nursing facility after TJA. This study served as proof of concept that ML could be used as a prediction tool not only for big data sets, but also for small data sets. Using ML techniques to predict length of stay and cost has led to monumental improvements in establishing value-based care in orthopaedic surgery.

In orthopaedic surgery, AI and ML-based techniques have demonstrated utility in predicting outcomes related to orthopaedic oncology, PROMs, length of stay, and cost. While ML techniques for survival in orthopaedic oncology have not yet been perfected, ML has proven to be effective with PROMs as well as predicting length of stay and cost. By using ML methods to make better outcome predictions, orthopaedic surgeons can improve their decision-making ability, which not only leads to better patient care, but also more efficient utilization of healthcare resources[52].

**IMAGING**

Since orthopaedic surgery diagnosis and treatment heavily rely on radiologic modalities [e.g., computed tomography (CT), magnetic resonance imaging (MRI), and conventional radiographs], the vast majority of AI and ML-based research has been applied to imaging. Recent advances in AI and ML have shown remarkable results with a few studies showing computers surpassing human test subjects at certain image interpretation tasks[53,54]. Within musculoskeletal medicine, DL has been shown to be useful for both text and image analysis[55-57]. ML and DL based techniques have the potential to assess earlier disease status and are currently the focus of significant orthopaedic research, particularly in the following subspecialties: Spine, joints/arthritis, trauma, and oncology (Table 3).

**Spine**

In spine surgery, technology has risen with the use of computer assisted navigation, robotic surgery, and augmented reality, all of which require reconstructions of the spinal column from CT or MRI scans[58-61]. This can only be achieved via ANNs and
Joints/arthritis

Annotated MRIs with information labels for each spine vertebrae used to accurately detect (99.8%) and label (97%) cervical and lumbar vertebrae.

Forsberg et al.[62] Spine Annotated MRIs with information labels for each spine vertebrae used to accurately detect (99.8%) and label (97%) cervical and lumbar vertebrae.

Hetherington et al. [64] Spine CNN successfully identified lumbar vertebral levels on ultrasound images of the sacrum.

Jamaludin et al.[65] Spine CNN model achieved 95.6% accuracy comparable to experienced radiologists in disc detection and labeling of T2 weighted sagittal lumbar MRIs.

Pesteie et al.[63] Spine Used ML system to detect laminae and facet joints in ultrasound images to assist in epidural steroid injection and facet joint injection administration.

Ashinsky et al.[71] Joints/arthritis ML algorithm predicted clinically symptomatic OA on T2 weighted maps of central medial femoral condyle with 75% accuracy.

Liu et al.[72] Joints/arthritis CNN performed rapid and accurate cartilage and bone segmentation within the knee joint.

Shah et al.[73] Joints/arthritis CNN used to automate the segmentation and measurement of cartilage thickness based on MRIs of healthy knees.

Xue et al.[70] Joints/arthritis CNN model trained to diagnose hip OA comparable to an attending physician with 10 years of experience in diagnosing hip OA.


Olczak et al.[74] Trauma DL networks identified fracture, laterality, body part, and exam view on orthopaedic trauma radiographs of the hand, wrist, and ankle.

Oh et al.[78] Oncology ML showed superior predictive accuracy in predicting pathological femoral fractures in metastatic lung cancer.

**Table 3 Summary of machine learning for orthopaedic surgery imaging applications**

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Subspecialty</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Helo et al.[66]</td>
<td>Spine</td>
<td>Neural network (93.2% accurate) and k-means approach (98% accurate) used on CT scans for segmentation and prediction of lumbar wedge fractures</td>
</tr>
<tr>
<td>Forsberg et al.[62]</td>
<td>Spine</td>
<td>Annotated MRIs with information labels for each spine vertebrae used to accurately detect (99.8%) and label (97%) cervical and lumbar vertebrae</td>
</tr>
<tr>
<td>Hetherington et al. [64]</td>
<td>Spine</td>
<td>CNN successfully identified lumbar vertebral levels on ultrasound images of the sacrum</td>
</tr>
<tr>
<td>Jamaludin et al.[65]</td>
<td>Spine</td>
<td>CNN model achieved 95.6% accuracy comparable to experienced radiologists in disc detection and labeling of T2 weighted sagittal lumbar MRIs</td>
</tr>
<tr>
<td>Pesteie et al.[63]</td>
<td>Spine</td>
<td>Used ML system to detect laminae and facet joints in ultrasound images to assist in epidural steroid injection and facet joint injection administration</td>
</tr>
<tr>
<td>Ashinsky et al.[71]</td>
<td>Joints/arthritis</td>
<td>ML algorithm predicted clinically symptomatic OA on T2 weighted maps of central medial femoral condyle with 75% accuracy</td>
</tr>
<tr>
<td>Liu et al.[72]</td>
<td>Joints/arthritis</td>
<td>CNN performed rapid and accurate cartilage and bone segmentation within the knee joint</td>
</tr>
<tr>
<td>Shah et al.[73]</td>
<td>Joints/arthritis</td>
<td>CNN used to automate the segmentation and measurement of cartilage thickness based on MRIs of healthy knees</td>
</tr>
<tr>
<td>Xue et al.[70]</td>
<td>Joints/arthritis</td>
<td>CNN model trained to diagnose hip OA comparable to an attending physician with 10 years of experience in diagnosing hip OA</td>
</tr>
<tr>
<td>Krusza et al.[75]</td>
<td>Trauma</td>
<td>ML improved hip fracture detection beyond logistic regression using dual x-ray absorptiometry</td>
</tr>
<tr>
<td>Olczak et al.[74]</td>
<td>Trauma</td>
<td>DL networks identified fracture, laterality, body part, and exam view on orthopaedic trauma radiographs of the hand, wrist, and ankle</td>
</tr>
<tr>
<td>Oh et al.[78]</td>
<td>Oncology</td>
<td>ML showed superior predictive accuracy in predicting pathological femoral fractures in metastatic lung cancer</td>
</tr>
</tbody>
</table>


DL through automated segmentation and detection of vertebrae. Numerous studies in the orthopaedic spine literature have analyzed the accuracy of DL techniques, especially for labeling and detection. For example, in Forsberg et al.[62], annotated MRIs with information labels for each spine vertebrae were used to detect and label cervical and lumbar vertebrae. The highest performance showed an accuracy of 99.8% for detection and 97% for labeling. Furthermore, Pesteie et al.[63] and Hetherington et al.[64] used ANNs trained with ultrasound images to automatically detect optimal vertebra level and injection plane for percutaneous spinal needle injections; Pesteie et al.[63] showed highest accuracy to be 95% and maximum precision to be 97%. ML and DL techniques have been shown to be useful in diagnosis as well. Jamaludin et al.[65] used DL techniques to read T2 weighted sagittal lumbar MRI images, automate the identification of disc spaces, grade the degenerative changes such as spondylolisthesis and central canal stenosis, and compare them to what experienced radiologists would do (Figure 3). The DL model performed almost as well as experienced radiologists on test data with a best accuracy rate of 95% for the prediction of spondylolisthesis. Because this model did not require labeling and feature description, authors believe the model will gain more accuracy and reliability with the addition of coronal and axial views. Al-Helo et al.[66] used ML techniques, specifically neural network and k-means approach, to learn lumbar wedge fracture diagnoses from CT image labeling for segmentation and prediction. The neural network showed an accuracy of 93.2% for lumbar fracture detection, while the k-means clustering approach attained an accuracy of 98%. These studies prove that the automation of radiologic grading is now on par with human performance; this can be incredibly beneficial in aiding clinical diagnoses in terms of grading and speed of analysis.

**Joints/arthritis**

OA, a highly prevalent disease associated with articular cartilage degeneration, can be effectively diagnosed in a cost-effective manner with X-ray imaging and in a more sensitive manner with MRI which can detect subtle morphologic changes in articular cartilage[67-69]. Throughout orthopaedic literature, DL has been used for hip and knee diagnostic purposes based on medical images. For the hip, Xue et al.[70] trained a CNN with 420 hip X-ray images to highlight saliency regions. These saliency regions allow the deep learning model to extract the necessary information in order to diagnose hip OA (Figure 4). The CNN model was able to achieve an accuracy of 92.8%, comparable...
Lalehzarian SP et al. ML in orthopaedic surgery

Figure 3 Input processing pipeline of T2 sagittal magnetic resonance imaging and output predictions of radiological features[65]. Reused with permission. Citation: Jamaludin A, Lootus M, Kadir T, Zisserman A, Urban J, Battié MC, Fairbank J, McCall I; Genodisc Consortium. ISSLS PRIZE IN BIOENGINEERING SCIENCE 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. Eur Spine J 2017; 26: 1374-1383.

Figure 4 Saliency images from left hip joint (1-5), right hip joint (6-8), and both hip joints (9,10)[70]. Reused with permission. Citation: Xue Y, Zhang R, Deng Y, Chen K, Jiang T. A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. PLoS One 2017; 12: e0178992.

to an attending physician with ten years of experience in diagnosing hip OA. For the knee, Ashinsky et al[71] used a ML algorithm on T2 weighted maps of the central medial femoral condyle in order to predict progression to clinically symptomatic OA; the ML algorithm was able to predict the onset of OA with 75% accuracy. Liu et al[72] applied a CNN to a knee image data set for bone and cartilage segmentation and labeling (Figure 5). Authors reported a performance accuracy of 75.3% for femoral cartilage labeling and 78.1% for patellar cartilage labeling. Similar to Liu et al[72], Shah et al[73] used a CNN to successfully automate cartilage segmentation methods and measurement of articular cartilage thickness. This study showed that ML can be used to analyze cartilage thickness in an automated and efficient manner. The results of the studies summarized above indicate that DL has promising potential in the field of intelligent medical image diagnosis practice, especially for hip and knee OA.

Trauma

For orthopaedic trauma, ML derived tools can be used on imaging techniques to assist in diagnostic ability, particularly for detection of fractures. Olczak et al[74] applied ML to 256000 orthopaedic trauma radiographs with good results compared to radiologists. In this study, a database of hand, wrist, and ankle radiographs were used and four outcomes - laterality, exam view, fracture, and body part - were identified (Figure 6). Five DL networks were used and reached 99% accuracy when identifying body part, 90% on laterality, 95% on exam view, and 83% on detecting fractures[74]. Furthermore, Kruse et al[75] used ML to predict hip fractures from dual x-ray absorptiometry; they found that ML could improve hip fracture prediction beyond logistic regression. In
orthopaedic trauma, ML based techniques have immense utility in predicting fractures. AI and ML based methods could be beneficial in the future of orthopaedic trauma as they may enhance workflow in the emergency department[76].

**Oncology**
In orthopaedic oncology, management of metastatic bone disease is a major focus, especially with respect to fracture and impending fracture care[77]. Oh et al[78] used ML on CT imaging and clinical features to extract radiologic features and derive predictions for pathological femoral fractures in metastatic lung cancer and compared the ML model with one that used CT features alone. The ML model, which included clinical features, showed superior predictive accuracy compared to the model that used CT features alone. By using AI and ML to accurately predict impending skeletal-related events, such as pathologic fracture, orthopaedic surgeons can prophylactically treat patients and thus improve patient outcomes[77].

**BASIC SCIENCE APPLICATIONS**
In the past, ML has been applied to basic science topics in medicine to predict chemical properties of drugs and proteins, predict vaccine immunogenicity, and identify promising drug targets[79-82]. In orthopaedic surgery, AI and ML has been applied to
Table 4 Summary of machine learning for orthopaedic surgery basic science applications

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Application</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begg et al[83]</td>
<td>Gait analysis</td>
<td>Used SVM to automate recognition of gait changes due to aging</td>
</tr>
<tr>
<td>Joyseeree et al [84]</td>
<td>Gait analysis</td>
<td>Used random forest, boosting, and SVM to identify disease on gait analysis data</td>
</tr>
<tr>
<td>Sikka et al[85]</td>
<td>Wearable technology</td>
<td>Utilized ML analytics via wearable technology to improve sports performance and identify risk factors for injury in sports</td>
</tr>
<tr>
<td>Cilla et al[86]</td>
<td>Implant design</td>
<td>ML techniques used to optimize short stem hip prosthesis to reduce stress shielding effects and achieve better short-stemmed implant performance</td>
</tr>
</tbody>
</table>

ML: Machine learning; SVM: Support vector machine.

more translational basic science concepts such as kinetics and gait analysis, wearable technology, and implant design[83-86] (Table 4).

It is well established in orthopaedic literature that aging influences various gait measures, such as gait velocity, stride length, and stance and swing phase times[87]. By applying ML to automate recognition of gait pattern changes, researchers can identify key variables of gait degeneration that might be predictors of falling behavior. Begg et al[83] used SVM, a specific ML approach, to automate recognition of gait changes due to aging using three types of gait measures: basic temporal/spatial, kinetic, and kinematic. When comparing gaits of twelve young participants to twelve elderly participants, the ML technique showed an overall accuracy of 91.7%. Furthermore, gait recognition improved when features were selected from different gait data types with an effective potential of 100% accuracy. Similarly, Joyseeree et al [84] applied ML algorithms, specifically random forest, boosting, and SVM, to gait analysis data for disease identification. Following a training and testing period, random forest and SVM had an accuracy of 100%, while boosting had an accuracy of 96.4%.

Another basic science application where ML has shown great promise is wearable technology. With the increase in wearable and portable technology, the general public as well as professional athletes have the power to monitor basic human physical and physiologic function that can combine with health records for analysis. Sikka et al used ML analytics via wearable technology such as camera-based monitoring systems, heart rate monitoring devices, radio-frequency identification tracking systems, and accelerometers, to improve sports performance[85]. Additionally, the data collected can be used to identify risk factors for injury in sports and therefore can proactively prevent injuries and direct injury prevention programs[85]. In the future, healthcare providers could utilize this information not only to develop optimal training programs for elite athletes while minimizing risk of injury and loss of play time, but also to create more cost-efficient care that is individually tailored for the average patient.

While shape optimization algorithms, which are different from ML, have previously been used to assess stem performance, the potential to further optimize short stem implants using ML has only recently been addressed[88-91]. For example, Cilla et al [86] used ANNs and SVMs to analyze four parameters with the end goal being to optimize short stem implant design, specifically for THA, to produce optimal performance, lack of bone resorption, and reduced stress shielding (Figure 7). They found that implants should be designed with a small stem length and a reduced length of the surface in contact with the bone to reduce stress shielding. The optimization approach using ML techniques can offer new and innovative possibilities in the design of hip implants and more. These analyses can be used to design new prostheses as well as aid orthopaedic surgeons in decision-making when choosing the most adequate implant.

CONCLUSION

In recent years, ML has garnered interest across various medical specialties and has proven its utility in orthopaedic surgery. Some studies even show that developed and validated ML models are capable of outperforming human specialists. Similarly, in orthopaedic surgery, ML has been incredibly useful in spine pathology detection, prosthesis control, gait classification, OA detection, and fracture detection. These
results corroborate the information that computers can outperform physicians in numerous tasks, even in orthopaedics. By and large, ML has diagnostic and prognostic uses that with continued research can offer more implications regarding orthopaedic treatment. With its surging trend of interest, AI and ML is expected to see an increase in use with risk assessment, outcomes assessment, imaging, and basic science applications in orthopaedics. Furthermore, because ML provides physicians the unique opportunity to understand their patients better, physicians should be trained to use these methods effectively in order to improve orthopaedic patient care.

REFERENCES


32 Voskuil T, Hageman M, Ring D. Higher Charlson Comorbidity Index Scores are associated with...


48 Ramkumar PN, Muschler GF, Spindler KP, Harris JD, McCalloch PC, Mont MA. Open mHealth Architecture: A Primer for Tomorrow’s Orthopedic Surgeon and Introduction to Its Use in Lower Extremity Arthroplasty. *J Arthroplasty* 2017; 32: 1058-1062 [PMID: 27956125 DOI: 10.1016/j.arth.2016.11.019]


